Navy Personnel Research and Development Center

San Diego, CA 92152-6800 TR 89-14 July 1989

2

The Effect of Incentives on the Reliability and Validity of Cognitive Speed Tests

NPRDC TR 89-14 July 1989

The Effect of Incentives on the Reliability and Validity of Cognitive Speed Tests

Dennis P. Saccuzzo
San Diego State University
San Diego, California 92182

Gerald E. Larson
Personnel Systems Department
Navy Personnel Research and Development Center
San Diego, California 92152-6800

James Brown
San Diego State University
San Diego, California 92182

Reviewed by
Robert F. Morrison

Approved by
John J. Pass
Director, Personnel Systems Department

Released by
B. E. Bacon
Captain, U.S. Navy
Commanding Officer

and

J. S. McMichael
Technical Director

Approved for public release;
distribution is unlimited.

Navy Personnel Research and Development Center
San Diego, California 92152-6800

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC TR 89-14	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION San Diego State University; 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Navy Personnel Research and Development Center (Code 12)
6c ADDRESS (City, State, and ZIP Code) San Diego, CA	7b ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800
8a NAME OF FUNDING/SPONSORING ORGANIZATION Chief of Naval Research; 8b OFFICE SYMBOL (If applicable) ONT	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State and ZIP Code) Washington, DC 20350	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 62233N; PROJECT NO RM33M20; TASK NO 4a; WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

The Effect of Incentives on the Reliability and Validity of Cognitive Speed Tests

12 PERSONAL AUTHOR(S)

D. P. Saccuzzo, G. E. Larson, J. Brown

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical Report	FROM FY87 TO FY88	1989 July	48

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Cognitive speed, reaction time, inspection time, information processing, motivation, incentives, intelligence.
05	09		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

In the present study, financial incentives were used to motivate test takers, so that the effect of motivation on elementary cognitive tests could be determined. One hundred and nine male and female volunteer college students were tested on a battery of microcomputerized cognitive tests. One hundred of these subjects returned for a second session in which they were randomly assigned to an incentive or no incentive condition and then retested. The effort expended on the tests was measured via heart rate, skin conductance, and a self-report questionnaire pertaining to the perceived level of difficulty of the tests and amount of effort expended on them. Criterion measures, including the Advanced Otis-Lennon Test of Mental Abilities, Standard and Advanced Raven Progressive Matrices, and scores on the Scholastic Aptitude Test were also taken. The findings revealed that incentives led to better performance only on the most complex task in the study. In no case, however, did incentives affect the overall IQ-performance correlation for the tests used in the battery. These results support the view that correlations between cognitive speed and intelligence reflect common mental capacities, rather than some affective variable such as motivation.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL Gerald E. Larson	22b TELEPHONE (Include Area Code) (619) 553-7656; 22c OFFICE SYMBOL Code 12

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

FOREWORD

This report discusses the effect of motivation on certain cognitive speed tests being considered as new measures of mental ability. The ultimate goal of the research is to improve the predictive power of the Armed Services Vocational Aptitude Battery (ASVAB). The present work was undertaken out of concern that the "cognitive speed" dimension itself might simply reflect motivational differences among the test takers. If so, there would be no scientific justification for viewing cognitive speed as an aspect (or substrate) of intelligence. Our findings, however, suggest that cognitive speed per se is largely unrelated to motivation, and that the development and validation of these tests appears theoretically justified.

The work was conducted under the Personnel Performance Prediction (PPP) project (Work Unit No. 62233N RM33M20.03), which was sponsored by the Office of the Chief of Naval Research (Code 222) and the Office of the Assistant Secretary of Defense (Force Management and Planning /MN&PP).

B. E. BACON
Captain, U.S. Navy
Commanding Officer

J. S. McMICHAEL
Technical Director

SUMMARY

Problem

There is considerable evidence linking an individual's performance on cognitive speed tests (e.g., choice reaction time and inspection time) and general intellectual capacity. This linkage has led to interest in the armed forces regarding the possible use of cognitive speed measures in personnel assessment. Though studies to date have shown that such cognitive speed tests can be constructed with adequate test-retest reliability, and that, in some cases, predictive validity also appears promising, use of tests will likely remain controversial as long as there are gaps in demonstrating construct validity. More specifically, it has yet to be shown that performance on cognitive speed tests is not partly the result of an affective variable such as motivation.

Purpose

The purpose of the present investigation was to support the construct validity of cognitive speed tests by evaluating the role of incentives on performance and the relationship between arousal and performance. Two specific questions were addressed: (1) How does motivation, induced through incentives, affect task performance on cognitive speed tests? and (2) How do motivating conditions affect the IQ-performance correlation for cognitive speed tests?

Approach

A battery of three cognitive speed tests--Inspection Time (IT), a version of Posner's letter matching task (NIPI), and the Mental Counters Test (MCT)--were administered to 109 male and female volunteer college students. One hundred of these subjects returned for a second session, identical to the first except for the order of task presentation. Half of the subjects were randomly assigned to an incentive condition, in which they were offered up to $20.00 if they could improve their performance. For both testing sessions, heart rate and skin conductance were recorded prior to (baseline) and during task performance. Following each cognitive speed test, subjects also responded to a self-report questionnaire that asked them to rate the test in terms of difficulty and to indicate the amount of effort they expended. Finally, subjects were group tested on three major IQ tests: the Advanced Otis Lennon, the Advanced Raven Progressive Matrices, and the Standard Raven Progressive Matrices. Subjects' scores on the Scholastic Aptitude Test, as well as their high school and freshmen grade point averages, were recorded from their official university transcripts.

Results

The data were analyzed through a variety of procedures including T-tests, analysis of variance, and correlational analysis. Analysis of variance revealed that incentives affected performance primarily on the most difficult level of the Mental Counters Test. Incentives did not affect performance on the reaction time and inspection time tests. Incentives did have a substantial effect on subject self-reported effort; subjects in the incentive group reported that they tried harder in session 2. Other analyses revealed that there were no significant changes in the IQ-performance correlation, whether or not subjects had incentives.

Discussion and Conclusions

In general, the data reveal little effect of incentives on the performance of cognitive speed tests, with the exception of the fastest rate of presentation on the mental counters test. Moreover, incentives had little effect on the IQ-performance correlations for the tasks in the study. This finding supports the construct validity of cognitive speed tests by demonstrating that speed-IQ relationships are most likely the result of command demands on intellectual capacities rather than individual differences in motivation.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Within the last 5 years, considerable evidence has linked performance on a variety of cognitive speed indices to intelligence (Jensen, 1987a, 1987b, 1987c; Saccuzzo & Larson, 1987). Choice reaction time (Jensen, 1982), intra-individual standard deviation of reaction time (Barrett, Eysenck, & Lucking, 1986; Vernon, 1983; Vernon, Nador, & Kantor, 1985), inspection time (Brand & Deary, 1982; Nettlebeck & Kirby, 1983) and skill at various other tasks that require few intellectual demands have been found to correlate with conventional intelligence tests of a much more complex nature. Theoretically, such correlations emerge because cognitive speed tests measure a basic underlying capacity related to intelligence. A more specific hypothesis holds that mental speed is primarily related to a general factor of intelligences, or Spearman's g (Jensen, 1987c).

Working under the hypothesis that cognitive speed tests might actually enhance the military's assessment of general intelligence, the Navy Personnel Research and Development Center (NAVPERSRANDCEN) embarked on a series of investigations to evaluate these new measures. Studies thus far have primarily examined either the psychometric characteristics of cognitive speed tests (Saccuzzo & Larson, 1987; Larson, Merritt, & Lattin, 1988), and/or their predictive validity (Larson & Rimland, 1984). These studies indicate that cognitive speed tests can be constructed with adequate test-retest reliability, and that, in some cases, predictive validity also appears promising. There are still gaps, however, in demonstrating construct validity, which is the focus of the present research.

To demonstrate construct validity, one must show that a test is associated with variables to which it is theoretically related, and unassociated with theoretically distinct variables. The former requirement is referred to as convergent validity and the latter as discriminant validity (Campbell & Fiske, 1959; Cronbach & Meehl, 1955). The convergent validity of cognitive speed tests has been repeatedly demonstrated, by showing that such tests are correlated with each other and with psychometric tests of general intelligence, as noted above. Less is known about discriminant validity. Few studies, for instance, have shown that performance on cognitive speed tests is not simply the result of an affective variable such as motivation. The choice of this example is not arbitrary: consider that while more intelligent subjects have faster reaction times, so do subjects who are motivated through reinforcement or knowledge of results (Lawler, Obrist, & Lawler, 1976; Weinstein, 1981).

Hence, the construct validity (and theoretical justification) of cognitive speed tests is in doubt as long as it can be argued that brighter subjects do well on such tests simply because they approach such tasks with greater zeal, and not because of some inherent ability that underlies intelligent behavior. A basic unanswered question, then, raised by a number of psychologists (Keating & MacLean, 1987; Marr & Sternberg, 1987) concerns what is being measured by cognitive speed tests--motivation or basic capacity. The present investigation addressed this unanswered question through empirical study of two issues: (1) How does motivation, induced through incentives, affect task performance on cognitive speed tests that are of interest to the armed forces? In other words, to what extent does more effort produce better performance? and (2) How do motivating conditions affect the IQ-performance correlation?

As an additional aid in understanding the nature of cognitive speed tests, direct measurement of performance was, in the present study, supplemented by physiological

measurement of arousal during performance, and by a subjective measure taken immediately after each task in which subjects were asked for an estimate of the effort they had expended.

Two arousal measures were used, heart rate (HR) and skin conductance (SC). The use of these measures of arousal was based on the notion that as task demands increase or when more effort is expended, bodily systems may become activated (i.e., aroused) as resources are marshaled in the service of this increased effort (Gopher & Donchin, 1986; Kahneman, 1973). Theoretically, as task difficulty increases, a subject must expend more effort to maintain the same level of performance. If such resources are available, performance may remain unchanged. It is only when the task demands exceed a subject's actual capacity that performance will decline. Through the use of physiological measures, it may be possible to detect increased effort or capacity while performance remains constant. In addition, in employing physiological measures, it is possible to address two questions of interest in understanding cognitive speed tests: (1) Is there a relationship between performance on cognitive speed tests and physiological arousal? and (2) Is there a relationship between arousal, as indexed by heart rate and skin conductance, and intelligence, as indexed by traditional psychometric tests?

The use of subjective measures provided a second avenue for evaluating task difficulty and effort during performance. Though self-reports are limited--for example, they can only be based on those aspects of tasks performance of which a subject is consciously aware or chooses to report--they can provide useful information in conjunction with direct measures of performance and arousal. Subjective measures are relatively easy to obtain, have a high degree of face validity, and have been found to be extremely reliable (Gopher & Donchin, 1986) as well as valid for quantifying complex cognitive behavior (Geiselman, Woodward, & Beatty, 1982).

In the present investigation, motivation was manipulated through incentives for three types of cognitive speed tests. These were Inspection Time (IT), a choice reaction time letter matching test known as NIPI, and the Mental Counters (MC) Test. The effects of motivation were then analyzed in various ways, as reported below.

METHOD

Subjects

The subjects were 109 volunteer San Diego State University students from an introductory course in psychology who received course credit for their participation. They ranged in age from 17 to 37-years-old ($M = 19.24$, $SD = 3.24$). Sixty-five were female, 44 were male.

Because the ultimate goal of this research is to improve military testing, comparison of our student sample with samples of military recruits is appropriate. The dimensions of age, gender, and mental ability provide convenient benchmarks. The mean age of the students (19.24 years) is very similar to the mean age of 271 recruits (19.8 years) selected at random for a study at the Recruit Training Command (RTC), San Diego (see Larson, Merritt, & Lattin, 1988). Sixty-six percent of the present subjects were female, while none of the aforementioned RTC subjects, and only about 9 to 10 percent of the entire Navy enlisted force, are female. As we report below, however, no sex differences emerged on any of the experimental tasks, nor does the literature reveal sex differences for the types of variables included in the study. Finally, the present sample scored higher

on Ravens Progressive Matrices (a test of general intelligence) (mean = 22.6, SD = 5.02) than did the recruits tested by Larson et al. (mean = 18.7, SD = 5.6). Given the comparable ability variance among the students and recruits, however, there is no a priori reason to believe that correlational relationships should differ. In conclusion, while the present sample differs in some ways from military recruits, none of the differences suggest that the findings from one group should not apply to the other.

Procedure

Subjects were tested on a battery of microcomputerized cognitive speed tests, which were presented on an IBM-XT compatible computer with color monitor and standard keyboard. The order of test presentation was randomized for each subject, and each subject was individually tested under supervised conditions. Prior to and during each task, two physiological measures, HR and SC, were recorded. In addition, after completing each task, subjects completed a self-report questionnaire (SRQ) designed to elicit information concerning the perceived difficulty of the task and how much effort subjects expended on the task. Following completion of an entire session in which all tasks were administered, subjects were asked to return for a second, retest session. One hundred subjects returned for the retest and were randomly assigned either to an incentive or no incentive condition with the restriction that there be an equal number of males and females in each of the two conditions. Subjects in the no incentive condition were given the battery of tasks as in session 1, the only difference being a different random order of presentation of tasks. Subjects in the incentive condition were also given an identical battery, in a different random sequence, and were told that they could earn up to $20.00 to the extent that they improved their performance. A more detailed description of the cognitive speed tests, physiological measures, SRQ, incentive manipulation, and other details of the method follow.

1. Cognitive Speed Tests. The microcomputerized battery of cognitive speed tests consisted of IT, MC, and the letter matching test, NIPI.

a. Inspection Time (IT)

The IT task was a non-adaptive procedure based on the methods of Larson and Rimland (1984) and Saccuzzo and Larson (1987). In this task, a visual stimulus, known as the target or test stimulus, is briefly presented in the center of the cathode ray tube (CRT) screen. In the present study, the target consisted of two horizontal lines of unequal length, one 17.5mm, the second 14.3mm. The two lines appeared to the right or left of a central fixation point. The longer line appeared on the right or left on a random basis. Immediately following termination of the target, a backward visual noise mask was presented. The mask, known to limit the duration of the sensory signal delivered to the central nervous system (Felsten & Wasserman, 1980), consisted of a spatially overlapping line which completely superimposed over the target. Targets were presented at five different stimulus durations: 16.7, 33.4, 66.8, 100.2, and 150.3 msec., which corresponded to 1, 2, 4, 6, and 9 refresh cycles on the video monitor. There were 15 trials per stimulus duration, for a total of 75 trials. The various stimulus durations were presented in a completely randomized order. The subject's task was to make a forced-choice discrimination, indicating which of the two lines of the target was the longer, by pressing one of two keys on the microcomputer keyboard. The task began with a set of instructions, examples, and five practice trials, prior to the test proper. Subjects were given computer-generated visual feedback on their performance. The entire inspection time task was given first, second, or third, according to a prearranged random sequence.

b. Mental Counters (MC)

In the MC Test (Larson, 1986), subjects were asked to keep track of the values of three independent "counters," which changed rapidly and in random order. The task required subjects to simultaneously hold, revise, and store three counter values under severe time pressure. The counters themselves were represented as dashes on the video monitor (three side-by-side horizontal dashes in the center of the screen). The initial counter values were zero (0, 0, 0). When a small target (.25 inch, hollow box) appeared above one of the three dashes, the corresponding counter had to be adjusted by adding "1." When the target appeared below one of the three dashes, the corresponding counter had to be adjusted by subtracting "1" (see Figure 1). The test items varied both in the number of targets and the rate of presentation. In the present study, there were three different rates of presentation, one target every .167, .633, and 1.42 seconds, which were called fast, medium, and slow speeds, respectively. Order of presentation of speeds was either "fast/medium/slow" or "slow/medium/fast." Subjects who received the "fast/medium/slow" order for session 1 were given the reverse order for session 2, and vice versa. For each speed, there were 20 consecutive trials, half of which had five targets, half seven. Prior to the test proper, subjects were given instructions, examples, and practice to criterion (they had to obtain three consecutive correct responses). The maximum and minimum counter values used in the present study were +3 and -3, respectively. The task was to select, from among four choices, the correct list of final values for the three counters. Selection was made by pressing the proper key on the keyboard. Feedback was given only during practice, and not during the test proper. The entire MC Test was given first, second, or third according to a prearranged random sequence.

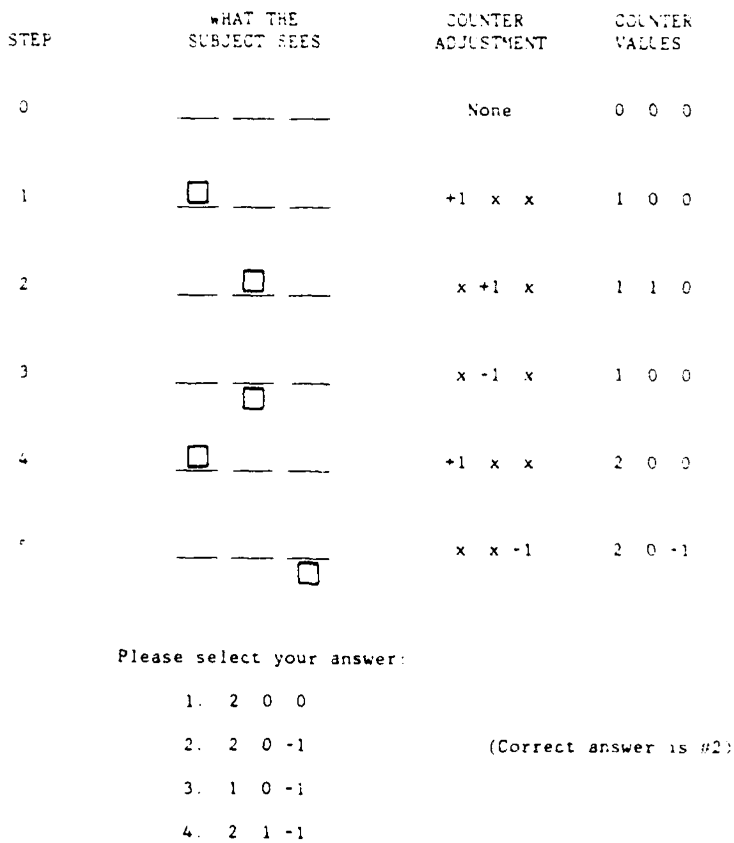

Figure 1. Sample item from mental counters test.

c. The Letter Matching Task (NIPI)

The NIPI was based on the work of Posner and Mitchell (1967). There were two subtests--Physical Identity (PI) Test and Name Identity (NI) Test. In the PI test, subjects were required to make judgments based on the physical appearance of two letters. For example, the letters 'a' and 'a' look the same, whereas, the letters 'A' and 'a' or 'g' and 'd' look different. Response times for same and different judgments were recorded for each trial. In the NI test, subjects were asked to respond on the basis of the names of two letters. For example, the letters 'a' and 'A' have the same name, while 'a' and 'c' do not.

On both tests, subjects were instructed to fixate on a period (".") located in the center of the screen. Following a random wait of 1.5 to 2.5 seconds, the period was replaced by two letters and the latency and accuracy of the subject's response were recorded. Reaction times greater than 2 seconds were discarded and new items presented to maintain a constant number of trials per subject. A count was kept of discarded trials. Each test consisted of undiscarded 34 trials. The PI test was always presented first. The entire NIPI task was presented first, second, or third according to a prearranged random sequence.

2. Physiological Measures. HR and SC were measured in one of two ways. For 20 subjects, HR was detected by biopotential electrodes placed on the right wrist and left ankle leading to a Beckman 511A Dynograph and type 98 57 cardiotachometer coupler calibrated between 30 and 120 bpm. SC for the same subjects was measured from two Ag/AgCl electrodes (using NaCl paste) placed on the back of the right hand at least 2 cm apart. The type 9844 SC coupler uses a .5 V, constant voltage circuit, and the amplifiers were calibrated to produce 0.5 umhos/mm. Both HR and SC records were scored to the nearest 0.5 mm by two independent readers, and disagreements were resolved by a third reader. Past research using this scoring method has yielded inter-reader agreement consistently greater than 90 percent. The remaining subjects were monitored with a PC-based physiological recording system (J&J Enterprises I-330 PC System). HR was measured from a photoplethysmograph transducer placed on the palmar surface of the distal segment of the left fifth finger connected to a P-401 photoplethysmograph module. The SC electrodes were connected to a GSR Model IG-3/T-68 module using a 15 Hz square wave, .3 V constant voltage circuit. For both recording systems, physiological activity was read every 10 seconds and averaged separately for the baseline period and each task.

3. Self-report Questionnaire (SRQ). A SRQ was administered following the entire IT task, each of the three levels of speed of the MC task, and each of the letter matching tasks (PI and NI). Subjects were asked to rank, on a scale from 1 to 6, each of the following questions:

a. How hard did you try? (A measure of effort or motivation.)

b. How difficult was the task? (A measure of task difficulty.)

c. How much better do you think you could have done had you used more effort? (A measure of unused effort or motivation.)

d. How much more effort could you have expended had the task been more difficult? (A measure of reserve effort.)

The actual questionnaire appears in Appendix A.

4. Incentive Manipulation. Subjects were randomly placed into an incentive or no incentive condition and retested at the same time of the day within a minimum of 2 days to a maximum of 2 weeks. Testing conditions were identical for both groups for the first session and for session 2 for the no incentive group except for the previously noted difference in randomizations and order of presentation in the MC Test. Testing conditions for session 2 were also identical for subjects in the incentive condition except that subjects in this group were offered incentives for better performance. Specifically, when they returned for session 2, subjects in the incentive group were told, "We will pay you to the extent that you can improve your performance. We will pay you up to $20.00 for improving your performance over the previous session. The more you improve, the more you will be paid up to $20.00." Because of the difficulties involved in calculating an immediate value for rate of improvement, and for human subjects purposes (we did not know if subjects could improve at all with incentives), all subjects in the incentive condition were paid $20.00 immediately after completing the battery regardless of improvement.

5. Criterion Measures. In addition to the above procedures, subjects were group tested on three IQ tests: The Raven Progressive Matrices, Advanced, Form 3 (Advanced Raven); The Raven Progressive Matrices, Standard (Standard Raven); and the Advanced Form of the Otis-Lennon Test of Mental Abilities (Advanced Otis-Lennon). The tests were administered in the following order on three separate occasions: Advanced Otis-Lennon, Advanced Raven, Standard Raven, with 40 minute time limits for each of the three tests, plus 10 minutes for practice on the Advanced Raven. In addition, subjects' high school and freshman grade point averages (GPAs), and scores on the Scholastic Aptitude Test (SAT) (SAT Verbal, SAT Math, SAT Total) were taken from their official transcripts. Subjects provided written informed consent to permit these measures to be taken from the Registrar's Office.

6. Summary of Variables. In sum, three major independent variables were examined: Groups (incentives versus no incentives), Sessions (session 1 versus session 2), and IQ. Each of these independent variables were evaluated as a function of three cognitive speed tests: IT, the three levels of speed of the MC Test, and the NIPI task. For each task, including each of the three levels of MC and each of the two matching tasks in the NIPI tasks, two physiological measures (HR and SC) and a subjective SRQ measure, consisting of four questions, were taken. For a summary of the variables and acronyms, see Table 1.

Table 1

Summary of Variables and Acronyms

I. Cognitive Tasks (Performance)

A. Inspection Time

ITTCA	Inspection Time, Total Correct, Session 1
ITTCB	Inspection Time, Total Correct, Session 2

B. Mental Counters

MCTSA	Mental Counters Test, Total, Slow Speed, Session 1
MCTSB	Mental Counters Test, Total, Slow Speed, Session 2
MCTMA	Mental Counters Test, Total, Medium Speed, Session 1
MCTMB	Mental Counters Test, Total, Medium Speed, Session 2
MCTFA	Mental Counters Test, Total, Fast Speed, Session 1
MCTFB	Mental Counters Test, Total, Fast Speed, Session 2
COUNTA	Counters Composite, Session 1
COUNTB	Counters Composite, Session 2

C. Letter Matching

PIMEDA	Physical Identity, Median Reaction Time, Session 1
PIMEDB	Physical Identity, Median Reaction Time, Session 2
PISDA	Physical Identity, Standard Deviation, Session 1
PISDB	Physical Identity, Standard Deviation, Session 2
NIMEDA	Naming Identity, Median Reaction Time, Session 1
NIMEDB	Naming Identity, Median Reaction Time, Session 2
NISDA	Naming Identity, Standard Deviation, Session 1
NISDB	Naming Identity, Standard Deviation, Session 2

II. Physiological Variables

A. Heart Rate

BASHRA	Baseline Heart Rate, Session 1
BASHRB	Baseline Heart Rate, Session 2
ITHRA	Inspection Time, Heart Rate, Session 1
ITHRB	Inspection Time, Heart Rate, Session 2
MCTSHRA	Mental Counters Test, Slow, Heart Rate, Session 1
MCTSHRB	Mental Counters Test, Slow, Heart Rate, Session 2
MCTMHRA	Mental Counters Test, Medium, Heart Rate, Session 1
MCTMHRB	Mental Counters Test, Medium, Heart Rate, Session 2
MCTFHRA	Mental Counters Test, Fast, Heart Rate, Session 1
MCTFHRB	Mental Counters Test, Fast, Heart Rate, Session 2
PIHRA	Physical Identity, Heart Rate, Session 1
PIHRB	Physical Identity, Heart Rate, Session 2
NIHRA	Naming Identity, Heart Rate, Session 1
NIHRB	Naming Identity, Heart Rate, Session 2

B. Skin Conductance

BASSCA	Baseline Skin Conductance, Session 1
BASSCB	Baseline Skin Conductance, Session 2
ITSCA	Inspection Time, Skin Conductance, Session 1
ITSCB	Inspection Time, Skin Conductance, Session 2
MCTSSCA	Mental Counters Test, Slow, Skin Conductance, Session 1
MCTSSCB	Mental Counters Test, Slow, Skin Conductance, Session 2
MCTMSCA	Mental Counters Test, Medium, Skin Conductance, Session 1
MCTMSCB	Mental Counters Test, Medium, Skin Conductance, Session 2
MCTFSCA	Mental Counters Test, Fast, Skin Conductance, Session 1
MCTFSCB	Mental Counters Test, Fast, Skin Conductance, Session 2
PISCA	Physical Identity, Skin Conductance, Session 1
PISCB	Physical Identity, Skin Conductance, Session 2
NISCA	Naming Identity, Skin Conductance, Session 1
NISCB	Naming Identity, Skin Conductance, Session 2

In the report, variable names not ending in "A" or "B" indicate that data from both sessions is being jointly described.

RESULTS

Overview

The results of the study are presented in four sections. In the first (the "Preliminary Analyses" below), we explore whether subgroups in our sample differ in ways that might bias later results. In the second part of the results, we describe the correlational relationships between the physiological variables, questionnaire responses, and test scores. The third part of the results section begins to address the main issue of the study (i.e., how effort and level of performance on cognitive speed tests are affected by incentives). Since this requires the use of group (i.e., incentive versus no incentive) comparison statistics, analysis of variance is the principal technique used in section three. In the fourth (and last) section of the results, we compare correlations between cognitive speed scores and intelligence for the incentive and no incentive groups, to determine whether motivation mediates the speed/intelligence relationship.

1. Preliminary Analyses

A preliminary analysis was conducted comparing the incentive and no incentive groups on their baseline performance levels for each of the tasks and on the IQ measures. The two groups did not differ on the Raven tests, the Otis, and SAT scores. Nor did they differ on any measures of performance. Thus, the two groups were not significantly different in their IQ and in their performance on any of the cognitive speed tests at session 1.

A second preliminary analysis was conducted comparing males and females at session 1 on their performance for each of the tasks in order to evaluate possible gender differences that might affect the subsequent analyses. No male-female differences were found for any of the tasks. Nor were there gender differences on the Advanced or Standard Raven. However, males ($M = 61$, $SD = 8.7$) had a significantly higher total correct score than females ($M = 56$, $SD = 9.4$) on the Advanced Otis (T (Pooled) $= 2.53$, $df = 98$, $P < .02$).

2. Correlational Relationships

The results described in the present section provide a background for our later discussion on incentive effects. Readers who are interested primarily in the effects of incentives may proceed to page 16.

Baseline Correlations

The analyses that follow are based on baseline (first session) correlations for the full sample. Table 2 shows the intercorrelations and basic statistics for the cognitive speed tests at session 1. Table 2 includes the following variables: IT, TOTAL CORRECT (ITTCA), MC, total correct, slow (MCTSA), medium (MCTMA), and fast (MCTFA), PI median reaction time in milliseconds (PIMEDA) and standard deviation (PISDA), and NI median reaction time (NIMEDA) and standard deviation (NISDA). As inspection of Table 2 reveals, there was a modest, but significant correlation between IT and MC medium and fast, large intercorrelations among the three MC tests, and large correlations among the various PI and NI measures.

Table 2

Intercorrelations and Basic Statistics for Cognitive Speed Tests: Session 1

Source	ITCA	MCTSA	MCTMA	MCTFA	PIMEDA	PISDA	NIMEDA	NISDA
ITCA	1.00	.05	.22*	.21*	.12	.08	.15	-.00
MCTSA		1.00	.46**	.51**	-.00	-.06	-.03	-.09
MCTMA			1.00	.45**	-.08	-.08	-.12	-.19*
MCTFA			1.00	.01	-.05	.07	-.11	
PIMEDA					1.00	.50**	.57**	.34**
PISDA						1.00	.15	.32**
NIMEDA							1.00	.50**
NISDA								1.00
Mean	57.18	15.51	17.48	13.15	561.67	140.37	682.85	152.29
S.D.	6.38	3.05	2.41	3.45	82.85	57.45	92.39	45.29
N	109	109	109	109	109	109	107	109

*P < .05.
**P < .01.

Table 3 shows the intercorrelations and basic statistics for HR during baseline and performance of the cognitive speed tests for session 1. Inspection of Table 3 reveals substantial intercorrelations, with a range of .74 to .89.

Table 3

Intercorrelations and Basic Statistics for Heart Rate During Baseline and Performance: Session 1

Source	BASEHRA	ITHRA	MCTSHRA	MCTMHRA	MCTFHRA	PIHRA	NIHRA
BASEHRA	1.00	.88**	.78**	.74**	.79**	.78**	.80**
ITHRA		1.00	.82**	.78**	.78**	.83**	.81**
MCTSHRA			1.00	.84**	.87**	.82**	.76**
MCTMHRA				1.00	.89**	.85**	.74**
MCTFHRA					1.00	.86**	.77**
PIHRA						1.00	.86**
NIHRA							1.00
Mean	83.77	82.42	84.28	84.97	84.36	82.69	83.5
S.D.	12.13	11.40	10.61	11.58	10.71	12.31	12.71
N	109	108	108	109	109	109	108

*P < .05.
**P < .01.

Table 4 shows the intercorrelations and basic statistics for SC during baseline and during performance of the cognitive speed tests at session 1. As with HR, the intercorrelations for SC during performance of the different cognitive tests are substantial.

Table 4

Intercorrelations and Basic Statistics for Skin Conductance During Baseline and Performance: Session 1

Source	BASSCA	ITSCA	MCTSSCA	MCTMSCA	MCTFSCA	PISCA	NISCA
BASSCA	1.00	.87**	.82**	.82**	.76**	.86**	.77**
ITSCA		1.00	.87**	.90**	.89**	.84**	.84**
MCTSSCA			1.00	.95**	.92**	.81**	.80**
MCTMSCA				1.00	.95**	.81**	.81**
MCTFSCA					1.00	.81**	.79**
PISCA						1.00	.95**
NISCA							1.00
Mean	2.79	3.14	3.02	3.22	3.05	2.87	2.92
S.D.	2.63	2.93	2.87	3.09	2.95	2.76	2.98
N	108	107	107	107	108	108	107

*P < .05.
**P < .01.

Test-retest Correlations

Table 5 shows the test-retest correlations, means, and standard deviations for the major variables in the study. (See Table 1 for a summary of acronyms). Inspection of Table 5 reveals that the test-retest coefficients for performance on the cognitive tests ranged from .77 for IT time to .37 for the standard deviation of performance during the physical identity (PI) task. In general, the test-retest coefficients are comparable to those reported by Saccuzzo and Larson (1987), in which no incentives were used. Baseline HR and baseline SC had reliabilites of .62 and .72, respectively. Test-retest correlations for HR during performance on the cognitive tasks ranged from .73 to .62. The test-retest correlations for SC during performance were somewhat lower, with a range from 0.67 to .38. Possible differences between the means on the two testings will be examined with appropriate statistical tests later in the results.

Physiological Variables as Predictors

The correlations of within-task HR and SC with performance on the cognitive speed tasks are shown in Table 6 for both sessions 1 and 2. Inspection of Table 6 reveals only four significant correlations, ranging from -.16 to -.27. By chance alone, we would expect only 1.6 correlations to be significant. Thus, a relationship between performance on the cognitive speed tests and the two physiological indices of arousal can be said to exist, but it is certainly weak. Furthermore, correlations between the two physiological

Table 5

Test-retest Correlations, Means, and Standard Deviations for Cognitive Speed Tests, Heart Rate, and Skin Conductance

Source	'R'xx	Mean[a]	S.D.[a]	n	Mean[b]	S.D.[b]	n
ITTC	.77*	57.18	6.38	109	58.51	6.52	100
MCTS	.59*	15.51	3.05	109	17.15	2.60	100
MCTM	.48*	17.48	2.41	109	18.14	1.88	100
MCTF	.44*	13.15	3.45	109	15.55	2.94	100
PIMED	.53*	561.67	82.85	109	549.11	78.54	100
PISD	.37*	140.37	57.45	109	113.53	34.38	100
NIMED	.70*	682.85	92.39	107	657.36	89.06	100
NISD	.54*	152.29	45.29	109	138.93	47.90	100
BASEHR	.62*	83.77	12.13	109	82.92	13.25	100
BASESC	.72*	2.79	2.63	108	2.60	2.01	99
ITHR	.62*	82.42	11.40	108	82.37	14.07	100
ITSC	.46*	3.14	2.93	107	2.87	2.76	100
MCTSHR	.71*	84.28	10.61	108	83.91	12.05	100
MCTSSC	.65*	3.02	2.87	107	2.84	2.85	100
MCTMHR	.71*	84.97	11.58	109	83.30	11.18	99
MCTMSC	.67*	3.22	3.09	107	3.00	2.90	100
MCTFHR	.71*	84.36	10.71	109	83.09	12.14	100
MCTFSC	.58*	3.05	2.95	108	2.82	2.92	100
PIHR	.73*	82.69	12.31	109	83.05	13.24	100
PISC	.38*	2.87	2.76	108	3.07	2.80	100
NIHR	.71*	83.50	12.71	108	83.33	12.63	100
NISC	.40*	2.92	2.98	107	3.06	2.87	100

*P < .01.

[a]First test session.

[b]Second test session.

Table 6

Correlations of Within-task Heart Rate and Skin Conductance with Performance for Sessions 1 and 2

Source	HR		SC	
	Session 1	Session 2	Session 1	Session 2
ITTC	.00	.04	.11	-.08
MCTS	-.08	.08	-.08	.10
MCTM	-.01	-.10	.11	.03
MCTF	-.16*	.09	.05	-.04
PIMED	-.01	.06	-.09	-.25**
PISD	.06	-.04	-.12	-.19*
NIMED	-.15	-.27**	-.15	-.05
NISD	-.04	-.10	-.13	-.04

*P < .05.
**P < .01.

indices themselves were virtually nonexistent; the highest observed was -.16 ($P < .05$) between HR and SC, recorded during the medium speed of mental counters. This is not surprising, since correlations between different physiological indices cited in the literature are typically low (Lacey & Lacey, 1974).

Table 7 shows the correlations between IQ and the physiological measures. See Table 1 for a description of the acronyms. Inspection of Table 7 reveals a few scattered, modest relationships, with a high of 0.32 ($P < .01$). The majority of the correlations failed to reach statistical significance, however, revealing a weak overall relationship between the physiological variables and IQ.

Questionnaire Correlations

Tables 8, 9, and 10 show the correlations between each of the four questions of the self-report questionnaire (SELFA1 = self-report question 1, session 1 . . . SELFB4 = self-report question 4, session 2) and performance of the cognitive speed tests, HR, and SC, respectively. Of particular interest is the relationship that emerged between question 2, "How difficult was the task?" and performance on the IT and MC tasks. These correlations reveal a significant relationship between perceived (reported) task difficulty and actual task performance. In both cases, subjects who thought the tests were easy performed better than subjects who felt the tests were relatively hard. Inspection of Table 9 reveals only a few low correlations between HR and self-report. Only one of the correlations between SC and the questionnaire, reported in Table 10, reached significance ($r = .15$, $P < .05$). Thus, just as the two physiological measures showed a weak interrelationship, so too there was a weak relationship between self-report and physiological indices.

Table 7

Correlations Between IQ and Physiological Measures

Source	ADVRAV	STRAV	SATV	SATQ	SATT	FRGPA	HSGPA	OTIS	Mean	S.D.	N
BASHRA	.22**	.10	.12	.18*	.19*	-.11	.09	.15	83.77	12.12	109
BASSCA	.08	.16	.11	.22*	.20*	-.15	-.04	.12	2.79	2.63	108
ITHRA	.21*	.14	.11	.19*	.19*	-.07	.09	.15	82.42	11.40	108
ITSCA	.11	.17	.18*	.16	.21*	-.02	-.11	.12	3.1	2.93	107
MCTSHRA	.18*	.16	.09	.10	.12	-.13	.02	.12	84.28	10.61	108
MCTSSCA	.02	.08	.19*	.17*	.21*	-.00	-.18*	.17*	3.02	2.87	107
MCTMHRA	.21*	.08	.15	.15	.18*	-.07	.13	.17*	84.97	11.58	109
MCTMSCA	.04	.10	.19*	.17*	.21*	-.06	-.18*	.19*	3.22	3.09	107
MCTFHRA	.19*	.16	.19*	.19*	.22**	-.07	.14	.26**	84.36	10.71	109
MCTFSCA	.03	.09	.18*	.15	.20*	-.07	-.01	.13	3.05	2.95	108
PIHRA	.22*	.14	.12	.15	.16	-.12	.15	.15	82.69	12.30	107
PISCA	.09	.12	.19*	.16	.21*	.00	-.01	.17	2.87	2.76	108
NIHRA	.27	.23*	.09	.17	.17	-.16	.06	.16	83.50	12.71	108
NISCA	.08	.09	.18*	.16	.20*	-.00	-.04	.16	2.92	2.98	107
BASHRB	.09	.02	.02	.06	.05	-.18*	-.04	.13	82.92	13.25	100
BASSCB	.02	.12	.05	.11	.10	.01	.23*	.14	2.60	2.01	99
ITHRB	.13	.12	.14	.06	.12	-.11	.03	.21*	82.37	14.07	100
ITSCB	.06	.18	.09	.14	.14	.00	-.09	.16	2.87	2.76	100
MCTSHRB	.19*	.11	.13	.09	.13	-.11	.01	.29**	83.91	12.05	101
MCTSSCB	.06	.06	.08	.21*	.18*	-.02	-.11	.19*	2.84	2.85	101
MCTMHRB	.14	.10	.16	.11	.16	-.14	.02	.29**	83.30	11.18	99
MCTMSCB	.06	.16	.04	.19	.14	.01	-.06	.14	3.00	2.90	100
MCTFHRB	.15	.15	.20*	.13	.20*	-.06	.06	.32**	83.09	12.14	100
MCTFSCB	.06	.17	.08	.21	.18	.06	-.01	.19*	2.82	2.92	100
PIHRB	.13	.13	.21*	.08	.17	-.08	.05	.28**	83.55	13.24	100
PISCB	.11	.22*	.10	.15	.15	-.00	-.06	.16	3.07	2.80	101
NIHRB	.11	.15	.17	.07	.14	-.05	.04	.24**	83.33	12.63	100
NISCB	.15	.23*	.18	.18	.22*	.04	-.01	.20*	3.06	2.87	101

*P < .05.
**P < .01.

Table 8

Correlations Between the Self-report Questionnaire and Performance on the Cognitive Speed Tests for Sessions 1 and 2

Source	ITTCA	ITTCB	Source	MCTCSA	MCTCSB	Source	MCTCMA	MCTCMB
SELFA1	-.11	-.15	SELFA1	-.03	.06	SELFA1	.02	-.02
SELFA2	-.40**	-.38**	SELFA2	-.22**	-.18*	SELFA2	-.26**	-.20*
SELFA3	-.08	-.17*	SELFA3	-.15	-.10	SELFA3	.01	-.10
SELFA4	-.01	.07	SELFA4	.08	.08	SELFA4	-.02	.06
Source	**MCTCFA**	**MCTCFB**	**Source**	**PIMEDA**	**PIMEDB**			
SELFA1	.01	.14	SELFA1	-.23**	-.07			
SELFA2	-.26**	-.08	SELFA2	-.02	-.01			
SELFA3	.07	.03	SELFA3	-.03	.06			
SELFA4	.06	.02	SELFA4	.09	.10			
Source	**PISDA**	**PISDB**	**Source**	**NIMEDA**	**NIMEDB**	**Source**	**NISDA**	**NISDB**
SELFA1	-.10	-.03	SELFA1	-.05	.06	SELFA1	.07	.04
SELFA2	.10	-.13	SELFA2	.12	.08	SELFA2	.11	-.02
SELFA3	.00	.19*	SELFA3	.04	.06	SELFA3	.01	.06
SELFA4	-.07	.08	SELFA4	-.00	.04	SELFA4	-.05	.08

*P < .05.
**P < .01.

Table 9

Correlations Between the Self-report Questionnaire and Heart Rate During Performance for Sessions 1 and 2

Source	ITHRA	ITHRB	Source	MCTSHRA	MCTSHRB	Source	MCTMHRA	MCTMHRB
SELFA1	-.17	-.12	SELFA1	.07	-.26	SELFA1	-.06	-.21*
SELFA2	-.08	.08	SELFA2	.05	-.15	SELFA2	.04	.03
SELFA3	.28**	.02	SELFA3	.18*	.01	SELFA3	.12	.17
SELFA4	.20*	.27**	SELFA4	-.00	-.06	SELFA4	.18*	.15

Source	MCTFHRA	MCTFHRB	Source	PIHRA	PIHRB	Source	NIHRA	NIHRB
SELFA1	-.01	-.11	SELFA1	-.05	-.08	SELFA1	-.03	-.21*
SELFA2	-.01	-.03	SELFA2	-.01	.00	SELFA2	-.01	-.06
SELFA3	-.03	.09	SELFA3	.01	-.17*	SELFA3	.17	-.08
SELFA4	.15	.09	SELFA4	.14	.06	SELFA4	.05	.08

*P < .05.
**P < .01.

Table 10

Correlations Between the Self-report Questionnaire and Skin Conductance During Performance for Sessions 1 and 2

Source	ITSCA	ITSCB	Source	MCTSSCA	MCTSSCB	Source	MCTMSCA	MCTMSCB
SELFA1	-.09	-.08	SELFA1	-.10	.07	SELFA1	-.13	.00
SELFA2	-.05	-.05	SELFA2	.00	-.00	SELFA2	-.03	.14
SELFA3	.08	-.01	SELFA3	.14	-.01	SELFA3	.08	-.04
SELFA4	.15*	-.03	SELFA4	.14	-.11	SELFA4	.08	-.04

Source	MCTFSCA	MCTFSCB	Source	PISCA	PISCB	Source	NISCA	NISCB
SELFA1	.03	.01	SELFA1	-.01	-.01	SELFA1	-.14	-.02
SELFA2	-.03	.07	SELFA2	-.09	.02	SELFA2	-.09	.08
SELFA3	.01	.02	SELFA3	-.11	.02	SELFA3	-.15	-.03
SELFA4	.05	-.11	SELFA4	.00	.05	SELFA4	-.01	-.10

*P < .05.
**P < .01.

Summary of Correlational Relationships. Correlations for the physiological variables were generally nonsignificant. For the questionnaire results, the most consistent finding was that subjects who thought that the MC and IT Tests were easy performed better than subjects who thought these tests were relatively hard.

3. ANOVAs: Comparison of Incentive versus No-incentive Groups

To determine the role of motivation in cognitive speed performance, we will now compare the results for the incentive and no-incentive groups. Group comparison analyses are broken into two sections: (1) The effect of incentive on level of task performance and (2) the effect of incentives on task-related effort and/or arousal, measured both physiologically (via HR and SC), and by a questionnaire. In the present study, level and effort are considered convergent indices for establishing an incentive effect.

Effects of Incentives on Level of Task Performance

Performance differences for IT were analyzed in a 2 (Group) X 5 (Stimulus Duration) X 2 (Sessions) repeated measures ANOVA with repeated measures on the last two factors. The group factor refers to the no-incentive vs. incentive groups. There was a significant main effect for Stimulus Duration, $F\ 4/93 = 271$, $P < .0001$, which is illustrated in Figure 2. As Figure 2 illustrates, the longer the exposure duration, the more accurately subjects' responded.

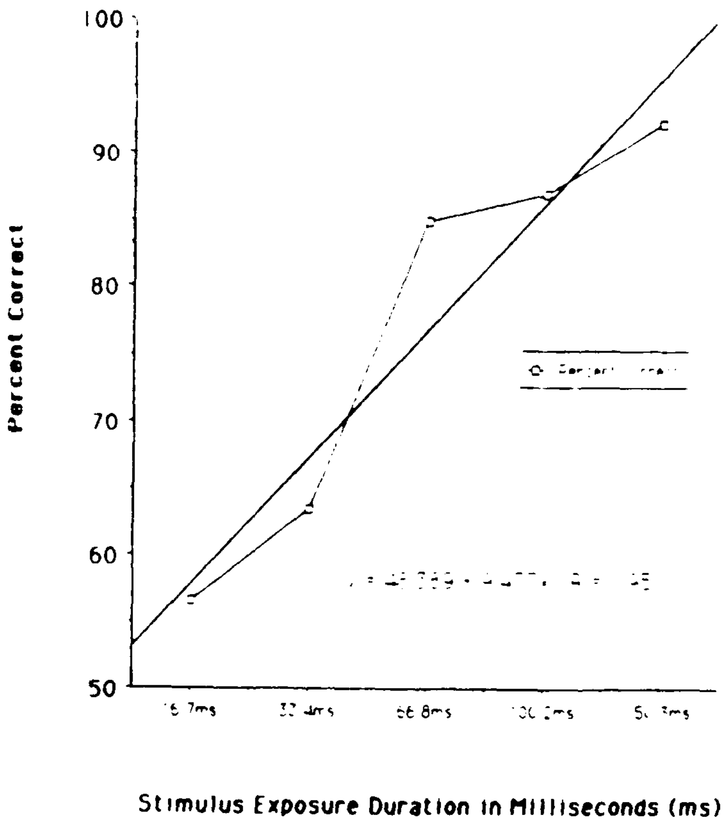

Figure 2. Main effect for stimulus duration.

In addition to the main effect for stimulus duration, there was a main effect for sessions, $F\ 1/96 = 9.96$, $P < .005$. The means for sessions 1 and 2 were 57.26 (75.88%) and 58.60 (77.76%), respectively, revealing a small, but statistically significant, increase in

performance at session 2. The Group by Sessions interaction failed to reach statistical significance, indicating that the sessions effect was the result of general improvement with practice rather than due to incentives. Thus, there was no effect of incentives on the global inspection time score.

Two significant interactions did emerge in the time analysis. There was a significant Stimulus Duration X Incentive interac 4/93 = 2.86, P < .03 and a significant Stimulus Duration X Session interaction, 2.94, P < .03. The Stimulus Duration X Incentive interaction is illustrated in Figure 3. As inspection of Figure 3 reveals, there were no differences between the groups at the shortest stimulus duration (where subjects were responding at, or just above, chance) and at the longest duration, where subjects were responding near the ceiling. Thus, the differences between the groups occurred at the middle stimulus durations, between chance and ceiling. It should be noted, however, that the effect shown in Figure 3 is summed across sessions. There was no Group X Sessions interaction, nor were there any significant triple interactions. The effects shown in Figure 3 reveal that the better performance of the incentive group was not due only to incentives, but to a more general tendency of this group to outperform the no incentive group with or without incentives. Finally, the Stimulus Duration X Sessions interaction, illustrated in Figure 4, shows superior performance at session 2 for all stimulus durations but the shortest where the reverse was found. Given that subjects were so close to chance at the shortest duration, the differences found at the 16.7 msec. duration can best be attributed to a chance fluctuation, with the clear overall trend of a practice effect independent of incentives.

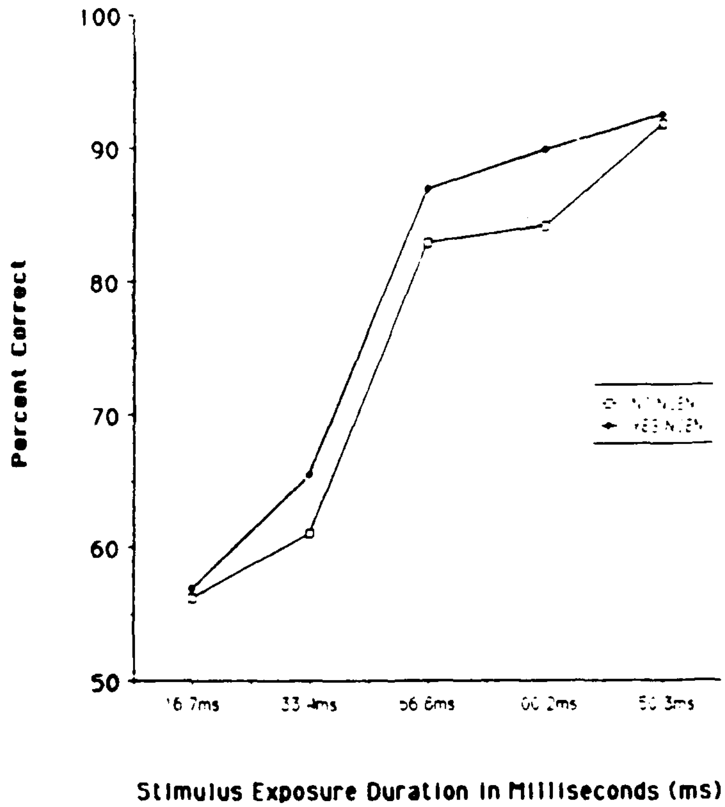

Figure 3. Stimulus duration x incentive interaction for inspection time.

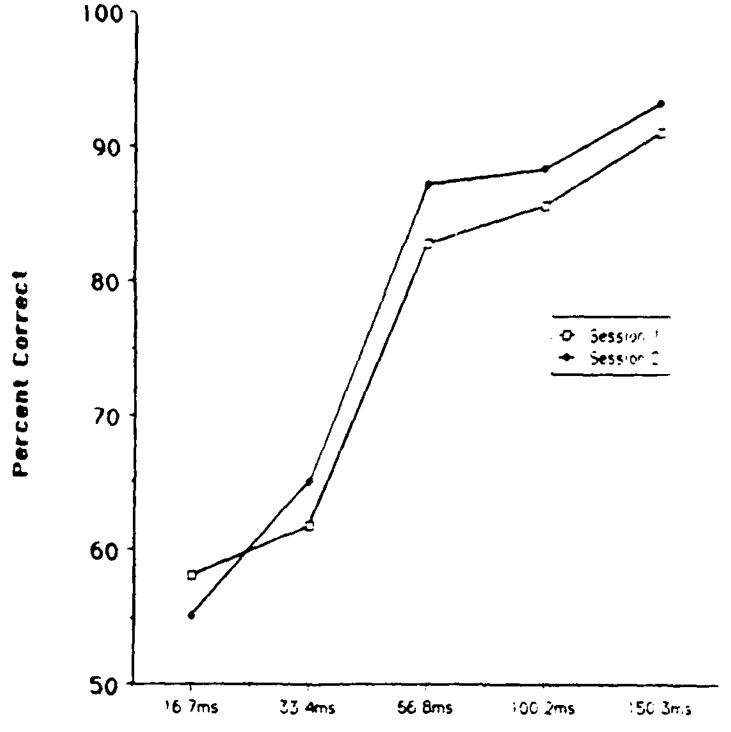

Figure 4. Stimulus duration x sessions interaction for inspection time.

To summarize the data on IT, incentives appear to boost the performance of subjects at the middle stimulus durations, but the effect was not powerful enough to cause a statistically significant difference in the overall score for the test.

Performance on the MC Test was analyzed in a 2 (Group) X 3 (Speed) X 2 (Sessions) repeated measures ANOVA with repeated measures on the last two factors. The total number correct at each speed served as the dependent measures. Significant were the main effects for speed, $F\ 2/94 = 130$, $P < .001$, and for sessions, $F\ 1/95 = 88$, $P < .001$. Means for the three speeds were 16.35, 17.80, and 14.36 for the slow, medium, and fast speeds, respectively. Newman-Keuls analysis (Winer, 1962) revealed that the differences between the slow and medium speeds were not statistically significant. However, subjects had significantly fewer correct responses for the fast speed when compared to both the medium ($P < .01$) and slow ($P < .05$) speeds. The main effect for sessions revealed an overall practice effect, with means of 15.38 and 16.96 for the first and second sessions, respectively.

In addition to the main effects, there were two two-way and one triple interaction effects. First, there was a significant Group X Sessions interaction, $F\ 1/95 = 7.70$, $P < .01$, which is illustrated in Figure 5. As Figure 5 shows, the two groups were roughly comparable at session 1 and both groups showed improvement at session 2. The incentive group, however, improved more, which is attributable to the incentives they received.

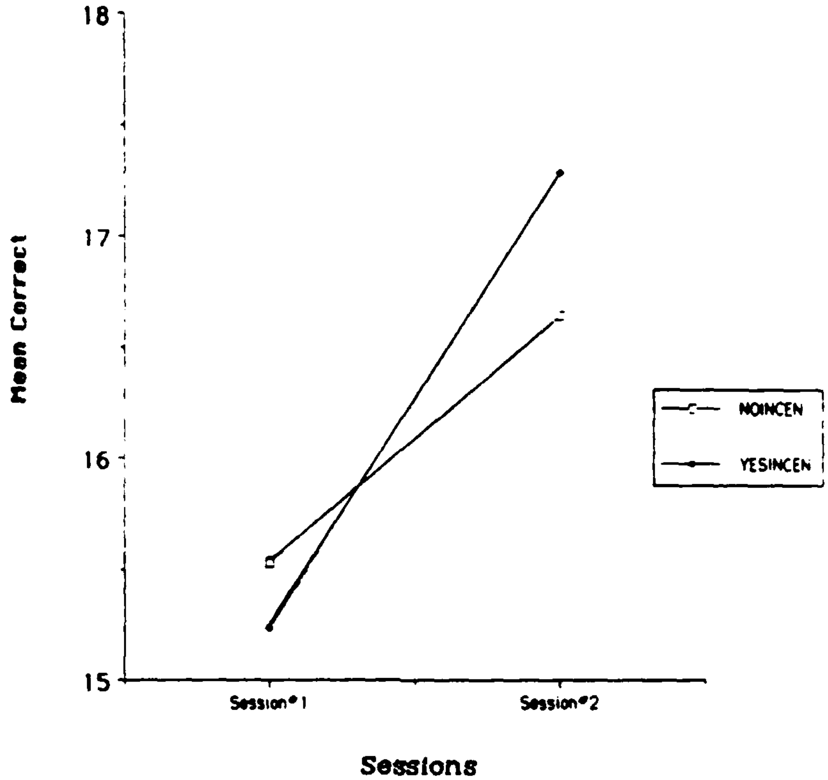

Figure 5. Group x Sessions interaction for mental counters.

The second two-way interaction was Speed X Sessions, $F\ 2/94 = 9.5$, $P < 0.001$ (see Figure 6). As inspection of Figure 6 reveals, this interaction effect was due to the greater levels of improvement found for the fast speed. Thus, improvement was greatest for the most difficult task in which subjects had the most room to improve.

Finally, analysis of performance for the MC Test revealed a significant Group X Speed X Sessions interaction, $F\ 2/94 = 4.0$, $P < .05$ (see Figure 7). This interaction indicated that the incentive group showed the greatest improvement at the fast speed. Thus, the effects of incentives were greatest on the most difficult task.

Differences in the reaction times for the letter matching task (NIPI) were analyzed in a 2 (Group) X 2 (Task-PI vs. NI) X 2 (Sessions) ANOVA with repeated measures on the last two factors. There were significant main effects for task, $F\ 1/95 = 422$, $P < .0001$, and for sessions, $F\ 1/95 = 14.95$, $P < 0.001$. The main effect for task revealed the common finding that reaction times are faster for physical identities (M = 552 msec.) than for name identities (M = 669 msec.). The main effect for sessions revealed the now familiar practice effect, with mean reaction times of 622 msec. and 600 msec. for sessions 1 and 2, respectively. There were no interaction effects for the letter matching task, and thus no effects of incentives.

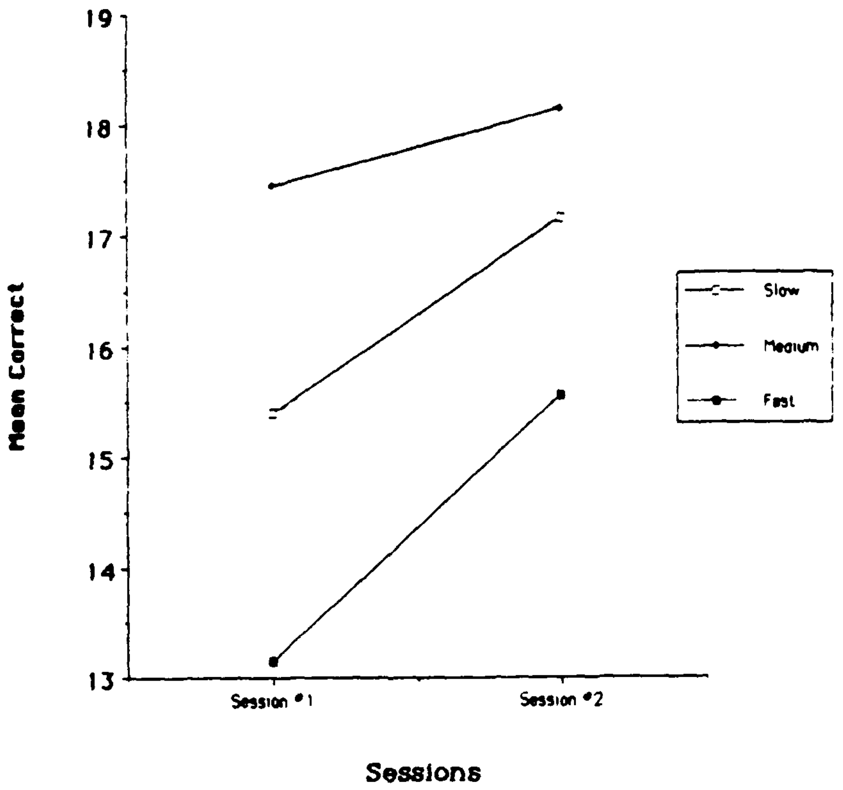

Figure 6. Speed x sessions interaction for mental counters.

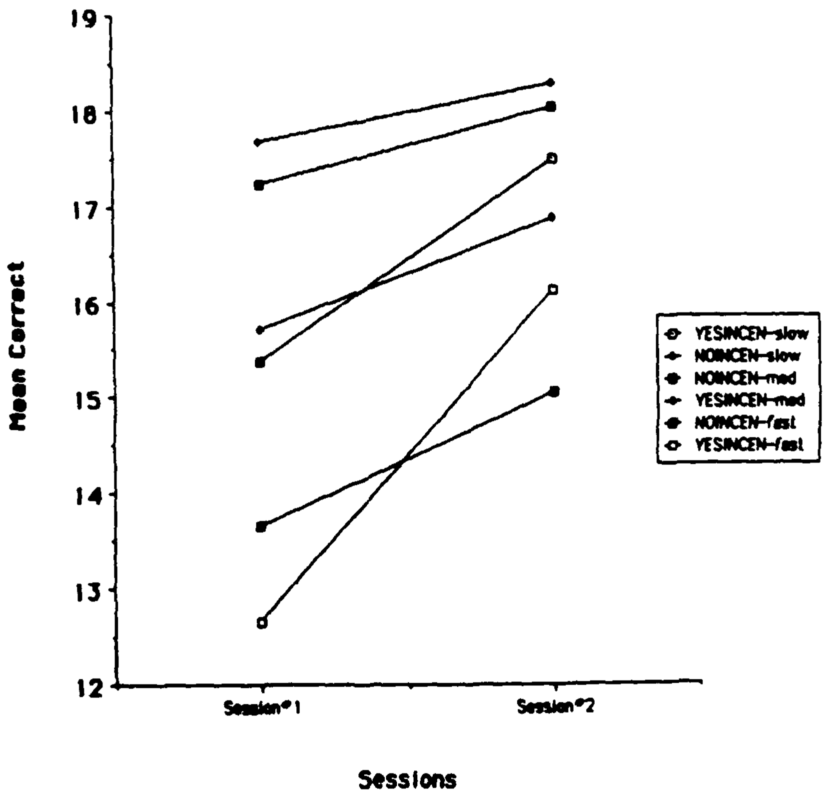

Figure 7. Group x speed x sessions interaction for mental counters.

Summary of Incentive Effects on Task Performance. In summary, only the MC Test showed a significant incentive effect on level of task performance. Incentives had no effect on overall performance for the IT and RT (NIPI) tasks. There was, however, a practice effect for all three cognitive speed tests; a speed effect for the MC Test; and a task effect for NIPI.

Effects of Incentives on Effort: Physiological Arousal

Two arousal measures were used, HR and SC. The use of these measures of arousal was based on the notion that as task demands increase or when more effort is expended, bodily systems may become activated (i.e., aroused) as resources are marshaled in the service of this increased effort (Gopher & Donchin, 1986; Kahneman, 1973). Thus, it is reasonable to expect that both task difficulty and incentive manipulations might have an effect on physiological arousal.

Heart Rate (HR) Analyses

HR data were analyzed in three different ways using: (1) difference scores as the dependent measure, (2) standard deviations as the dependent measure, (3) analysis of covariance with raw HR during the task as the dependent measure and baseline HR for session 1 as the covariate. (There were no significant group differences for baseline performance at sessions 1 and 2). **No significant incentive effect was found for any HR measure.**

The covariance analysis for IT, however, revealed main effects for sessions, $F\ 1/90 = 5.93$, $P < .02$, and for base HR, $F\ 1/90 = 168$, $P < .0001$. The main effect for sessions revealed that HR fell slightly from session 1 ($M = 82.41$) to session 2 ($M = 81.42$). Thus, there was a slight, but significant decrease in HR over sessions, which corresponded to the significant increase in performance. The significant base HR effect revealed that the incentive group ($M = 85.21$) had a significantly faster HR at session 1 than the no incentive group ($M = 81.96$), which reveals that the two groups showed arousal differences even before the incentive manipulation.

For the MC Test, the 2 (Group) X 3 (Speed) X 2 (Sessions) ANOVAs and covariance analysis yielded only two significant findings. There was a main effect for group for HR standard deviation, $F\ 1/82 = 8.22$, $P < .01$, and a main effect for base HR in the covariance analysis as reported above in the IT analysis. The main effect for group was due to the significantly higher HR variability, summed across all three counter speeds and sessions, in the incentive group ($M = 6.95$) compared to the no incentive group ($M = 4.46$). Thus, there were no incentive effects on HR; nor were there changes across sessions that corresponded to the performance changes that were found with MC.

For the 2 (Group) X 2 (Task) X 2 (Sessions) ANOVAs and covariance analysis for NIPI, there were no significant differences for difference scores (i.e., rise/fall over baseline). With HR standard deviation as the dependent variable, there was a main effect for group, $F\ 1/82 = 10.42$, $P < .002$. The incentive group ($M = 6.77$) showed considerably more variability, summed across both tasks and sessions, than the no incentive group ($M = 3.66$). Also significant was the main effect for sessions in the covariance analysis, $F\ 1/90 = 6.36$, $P < .02$, which revealed a slight decrease in the means for session 1 (82.71) compared to session 2 (82.55), when baseline performance at session 1 was used as a covariate.

In sum, while HR differences were found, they were not attributable to incentives.

Skin Conductance (SC) Analyses

SC data were analyzed in two different ways using: (1) difference scores as the dependent measure and (2) analysis of covariance with SC scores during the task as the dependent measure and baseline SC for session 1 as the covariate. (There were no significant group differences for baseline performance at sessions 1 and 2.)

For the IT task, the 2 (Group) X 2 (Sessions) repeated measures ANOVA revealed no statistically significant differences for the difference scores. The covariance procedure produced a main effect for sessions, $F\ 1/90 = 9.10$, $P < .004$, which revealed a slight decrease in SC between sessions 1 ($M = 2.86$) and 2 ($M = 2.83$), when baseline scores at session 1 are used as a covariate. This sessions effect paralleled the effects found for performance and HR. In addition, the covariance analysis produced a significant base SC effect, $F\ 1/90 = 94.79$, $P < 0.001$, which again revealed arousal differences between the incentive group ($M = 20.67$) and the no incentive group ($M = 2.44$) at session 1, before the incentive manipulation.

For the MC Test, the 2 (Group) X 3 (Speed) X 2 (Sessions) ANOVA for difference scores produced a main effect for Speed, $F\ 2/90 = 4.04$, $P < .03$. Mean SC scores for the slow, medium, and fast speeds, respectively, were -0.665, -1.303, and -0.859. The only significant difference, according to Newman-Keuls analysis, was between the slow and medium speeds.

The covariance analysis for MC revealed a main effect for sessions, $F\ 1/90 = 4.73$, $P < .04$. As with HR, there was a slight decrease in SC between session 1 ($M = 2.94$) and session 2 ($M = 2.89$), which paralleled the significant practice effect found in the analysis of performance. The covariance analysis also revealed a significant Group X Speed interaction effect, $F\ 2/89 = 3.22$, $P < .05$ (see Figure 8). Inspection of Figure 8 reveals that the effect for speed occurred primarily for the no incentive group between the slow and medium and slow and fast speeds. SC of the incentive group remained relative constant, and nonsignificantly different, across the three speeds. Finally, the covariance analysis revealed the significant base SC effect, which was reported in the IT analysis.

For the 2 (Group) X 2 (Task) X 2 (Sessions) ANOVA for NIPI, there were no significant differences when difference scores were used as the dependent measure. The covariance analysis showed only a significant sessions effect (in addition to the previously reported base SC effect), $F\ 1/90 = 10.76$, $P < .002$. Analysis of this session effect revealed that SC was lower for session 1 ($M = 2.66$) than it was for session 2 ($M = 3.10$).

In summary, there were few incentive effects on physiological indices of arousal.

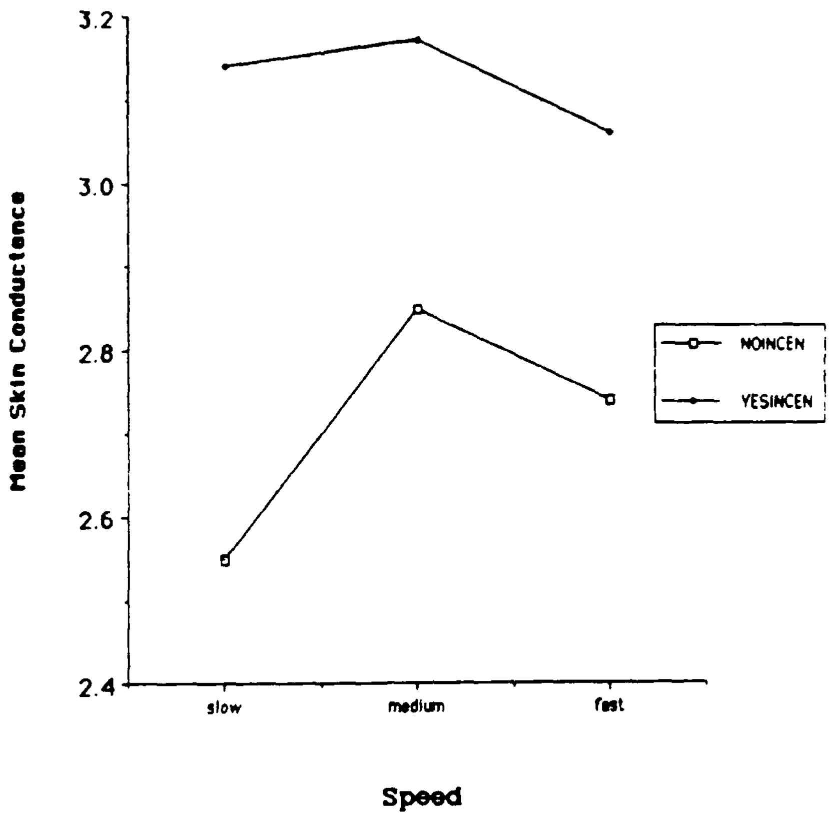

Figure 8. Group x speed interaction for skin conductance during mental counters.

Effects of Incentives on Effort: Self-report Questionnaire Responses

The questionnaire was designed to reveal the effect of incentives on perceived effort. Data for each of the four self-report questions were analyzed separately for each task. A host of significant effects were found.

For question 1, "How hard did you try?," the 2 (Group) X 2 (Sessions) repeated measures ANOVA for responses following IT resulted in a main effect for sessions, $F\ 1/96 = 8.13$, $P < .01$, and a Group X Sessions interaction, $F\ 1/96 = 6.17$, $P < .02$. The main effect for sessions revealed that overall, subjects said they tried harder on session 1 ($M = 4.44$) than on session 2 ($M = 4.12$). However, the sessions effect must be interpreted in light of the Group X Sessions interaction, which is shown in Figure 9. As Figure 9 shows, the incentive group reported that they tried about as hard on both sessions, and there was no significant difference in reported effort for this group. The no incentive group, by contrast, reported that they tried harder on session 1 than they did on session 2. Thus, the no incentive group showed a drop in reported effort, which the incentive group did not show.

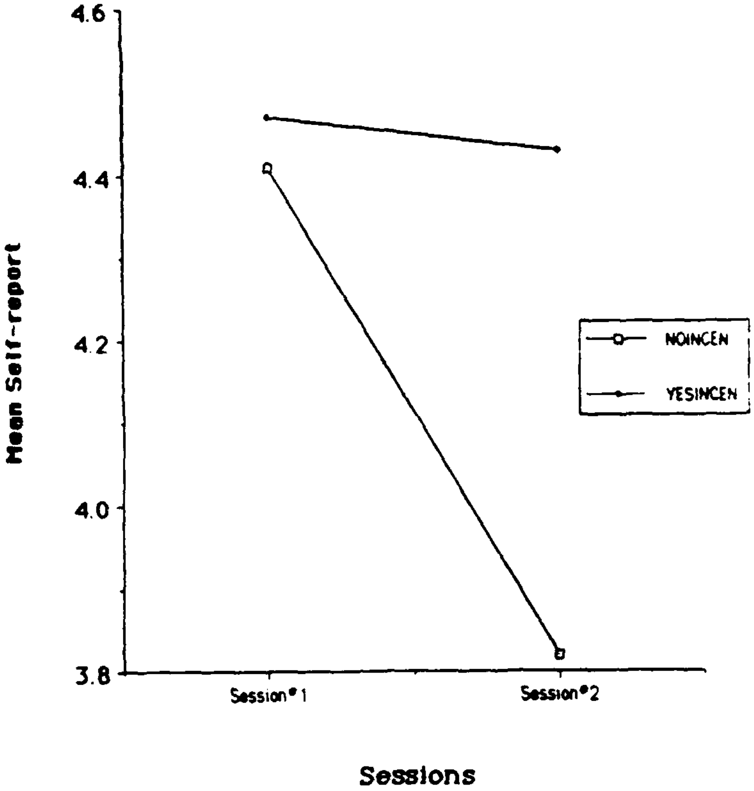

Figure 9. Group x sessions interaction for Question 1: Inspection time.

The 2 (Group) X 3 (Speed) X 2 (Sessions) repeated measures ANOVA for responses to question 1 following the MC Test revealed a significant main effect for speed, F 2/92 = 19.75, P < .0001 as well as a Group X Sessions interaction, F 1/98 = 10.68, P < .002. The main effect for speed revealed that subjects reported that the faster the speed, the harder they tried, with means of 3.80, 4.09, and 4.37 for the slow, medium, and fast speeds, respectively. Only the differences between the slow and fast speeds reached statistical significance according to Newman-Keuls analysis (P < .01). The Group X Sessions interaction is illustrated in Figure 10, inspection of which reveals that whereas the no incentive group said they tried less hard on session 2, the incentive group claimed just the opposite.

The 2 (Group) X 2 (Task) X 2 (Sessions) repeated measures ANOVA for NIPI also revealed a significant Group X Incentive interaction for question 1, F 1/99 = 6.25, P < .02. Figure 11 shows the same pattern as found for MC. Subjects in the no incentive group reported they tried less hard on session 2; the incentive group reported they tried harder on session 2.

In sum, for question 1, "How hard did you try?," there were significant Group X Session interactions for all three tasks. These interactions revealed that whereas the no incentive group reported expending less effort for session 2, the incentive group reported expending equal or more effort on session 2.

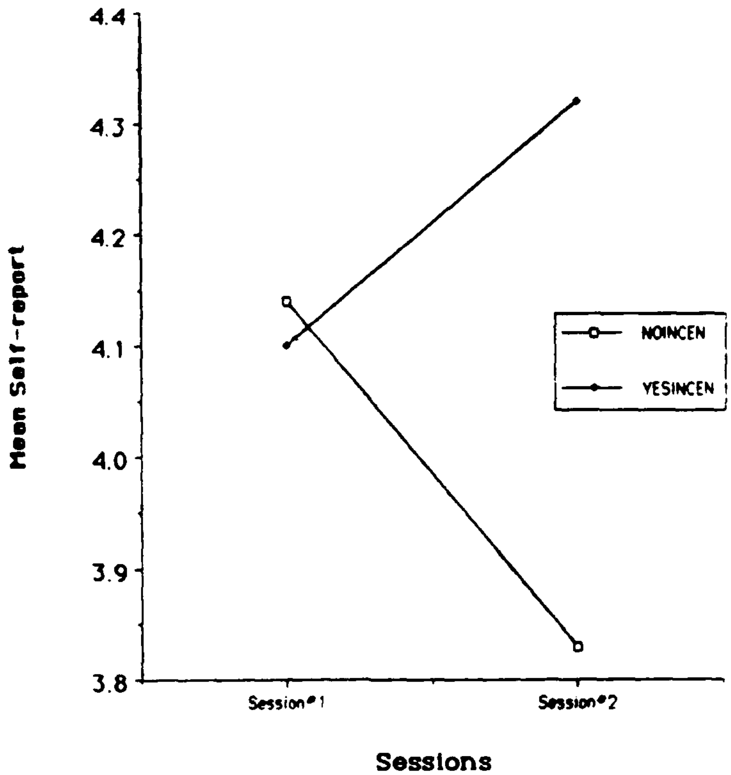

Figure 10. Group x sessions interaction for Question 1: Mental counters.

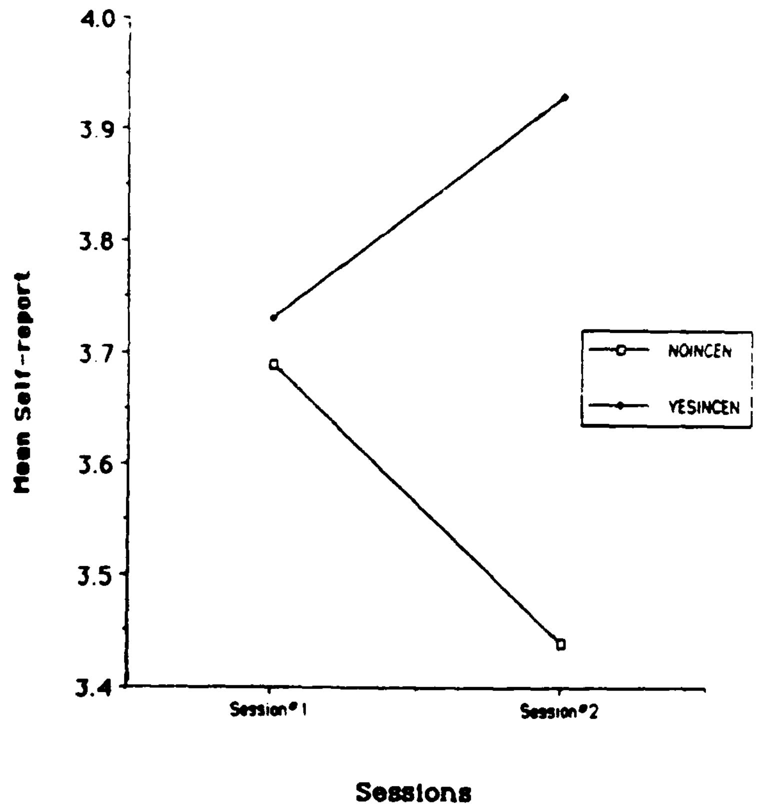

Figure 11. Group x sessions interaction for Question 1: Letter matching.

The results of the ANOVAs for question 2, "How difficult was the task?" are summarized in Table 11. Means for the three speeds in the MC Test were 3.09, 3.43, and 4.44 for the slow, medium, and fast speeds, respectively. Newman-Keuls analysis revealed a significant different ($P < .01$) only between the fast and medium speeds and between the fast and slow speeds. Thus, the fast speed clearly was perceived as the most difficult task. The main effect for sessions showed that overall, the MC Test, summed over speed, was seen as more difficult at session 1 ($M = 3.74$) than at session 2 ($M = 3.57$). The main effects for NIPI showed that the NI task was perceived as more difficult than the PI task and that, overall, the task was seen as easier on session 2.

Table 11

Summary of Results for Question 2

Inspection Time	2 (Group) X 2 (Sessions) repeated measures ANOVA. No significant differences.
Mental Counters Test	2 (Group) X 3 (Speed) X 2 (Sessions) repeated measures analyses. Main effect for Speed, $F\ 2/97 = 212$, $P < .001$. Main effect for Sessions, $F\ 1/98 = 4.88$, $P < .03$.
NPI	2 (Group) X 2 (Task) X 2 (Sessions) repeated measures ANOVA. Main effect for Task, $F\ 1/99 = 87.8$, $P < .001$. Main effect for Sessions, $F\ 1/99 = 4.08$, $P < .05$.

The results of the ANOVAs for question 3, "How much better do you think you could have done if you used more effort?" are summarized in Table 12. The Group X Sessions interaction for IT, illustrated in Figure 12, reveals that while the no incentive group believed that they could have done better at session 2, the incentive group reported just the opposite.

Table 12

Summary of Results for Question 3

Inspection Time	2 (Group) X 2 (Sessions) repeated measures ANOVA. Group X Sessions Interaction, $F\ 1/96 = 5.14$, $P < .03$.
Mental Counters Test	2 (Group) X 3 (Speed) X 2 (Sessions) repeated measures analyses. Group X Sessions Interaction, $F\ 1/98 = 12.47$, $P < .003$.
NIPI	2 (Group) X 2 (Task) X 2 (Sessions) repeated measures ANOVA. Task X Sessions Interaction, $F\ 1/98 = 6.27$, $P < .02$.

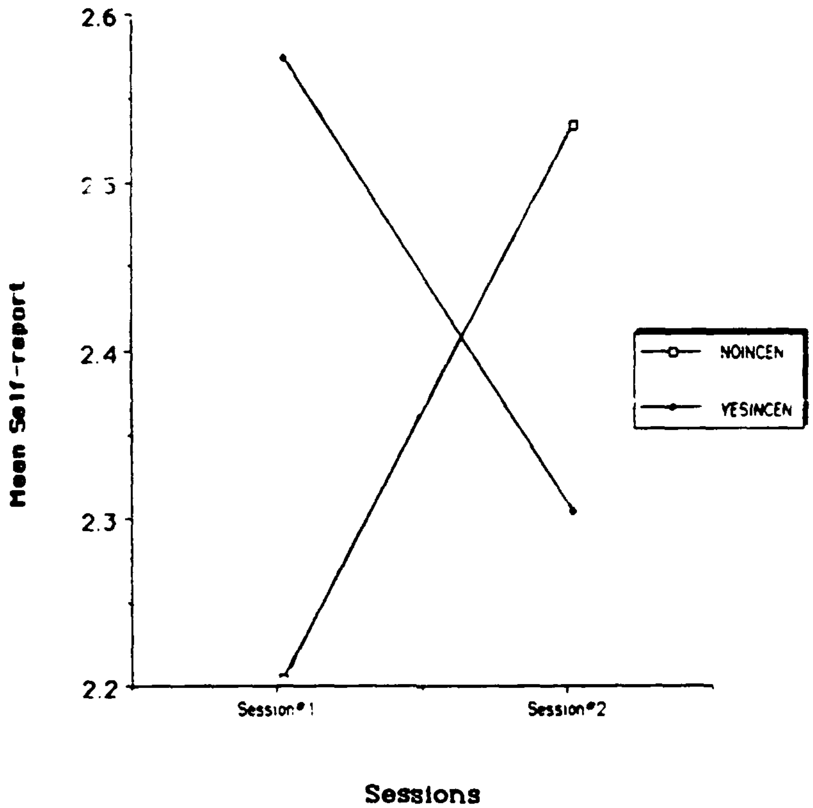

Figure 12. Group x sessions interaction for Question 3: Inspection time.

The Group X Sessions interaction for MC, illustrated in Figure 13, paralleled the Group X Sessions interaction found for IT. Whereas, the no incentive group thought they could have improved their performance, the incentive group apparently felt just the opposite.

Finally, the Task by Sessions interaction for NIPI is illustrated in Figure 14. This figure reveals that, overall, there was a significant decrease between sessions 1 and 2 for the PI task but not for the NI task.

The results of the ANOVAs for question 4, "How much more effort could you have expended had the task been more difficult?" are summarized in Table 13. The significant main effect in the IT analysis revealed that the incentive group reported that they could have used more effort had the task been more difficult, a finding paralleled by the main effect for group for the MC Test. The main effect for speed in MC reveals a parallel with question 2, and indicates that in general, the faster the speed, the more difficult the task. The triple interaction, illustrated in Figure 15, shows that for the fastest speed (i.e., the most difficult task) the no incentive group showed an increase in their mean response, whereas, the incentive group showed the opposite. Finally, the main effect for task in NIPI revealed a significantly higher mean for PI than for NI, again reflecting differences in perceived difficulty.

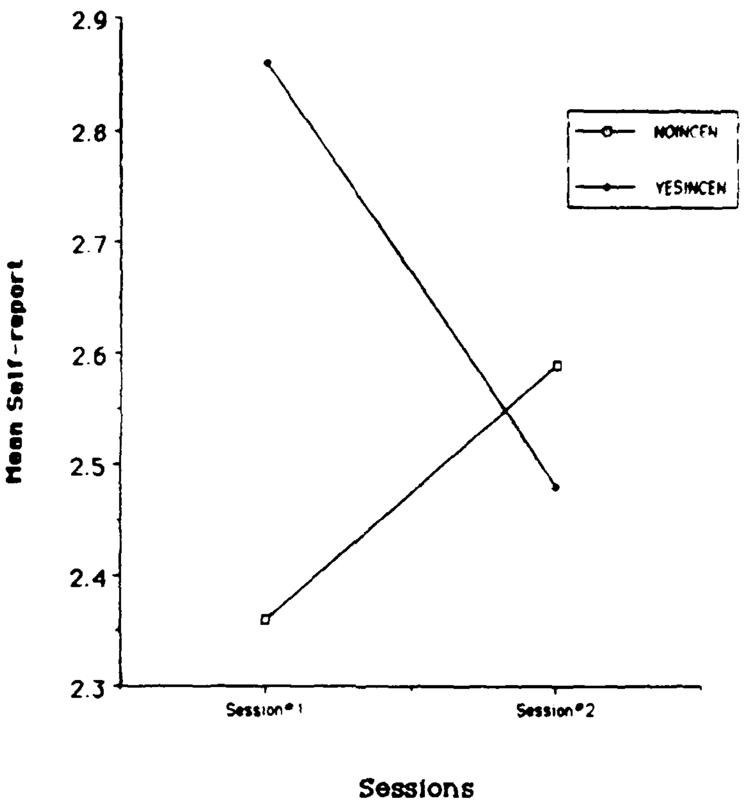

Figure 13. Group x sessions interaction for Question 3: Mental counters.

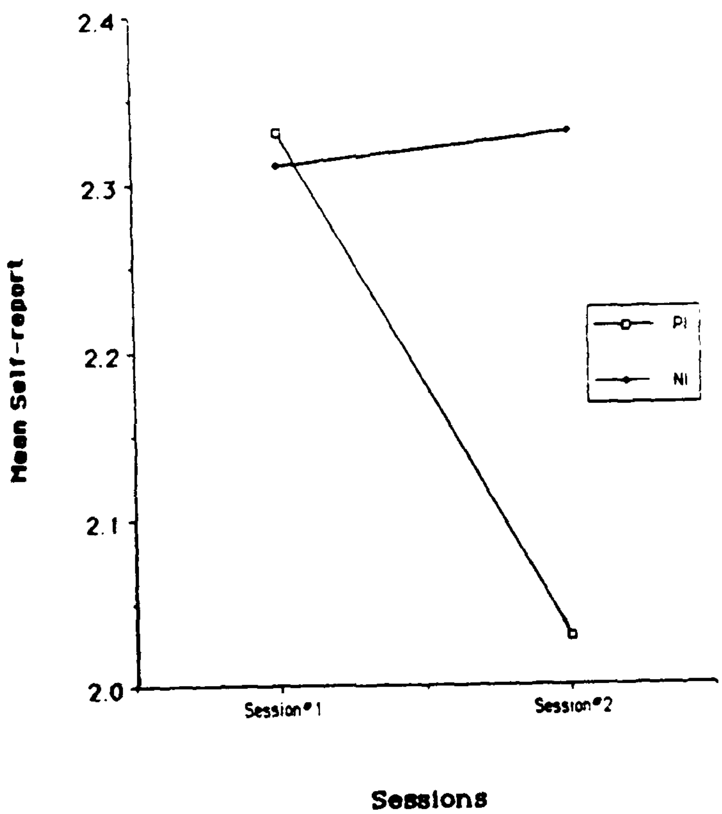

Figure 14. Task x sessions interaction for Question 3: Letter matching.

Table 13

Summary of Results for Question 4

Inspection Time	2 (Group) X 2 (Sessions) repeated measures ANOVA. Main effect for Group, $F\ 1/96 = 8.74$, $P < .004$. (Mean no incentive group = 2.46; mean incentive group = 3.0[illegible].)
Mental Counters Test	2 (Group) X 3 (Speed) X 2 (Sessions) repeated measures ANOVA. Main effect for Group, $F\ 1/98 = 5.80$, $P < .02$. (Mean no incentive group = 2.63; mean incentive group = 3.04.) Main effect for speed, $F\ 2/97 = 15.8$, $P < .0001$. (Mean slow = 3.09, Mean medium = 2.91, Mean fast = 2.58.) Group X Speed X Session Interaction, $F\ 2/97 = 4.36$, $P < .02$.
NIPI	2 (Group) X 2 (Task) X 2 (Sessions) repeated measures ANOVA. Main effect for Task, $F\ 1/99 = 166$, $P < .0001$. (Mean PI = 3.54, Mean NI = 3.23)

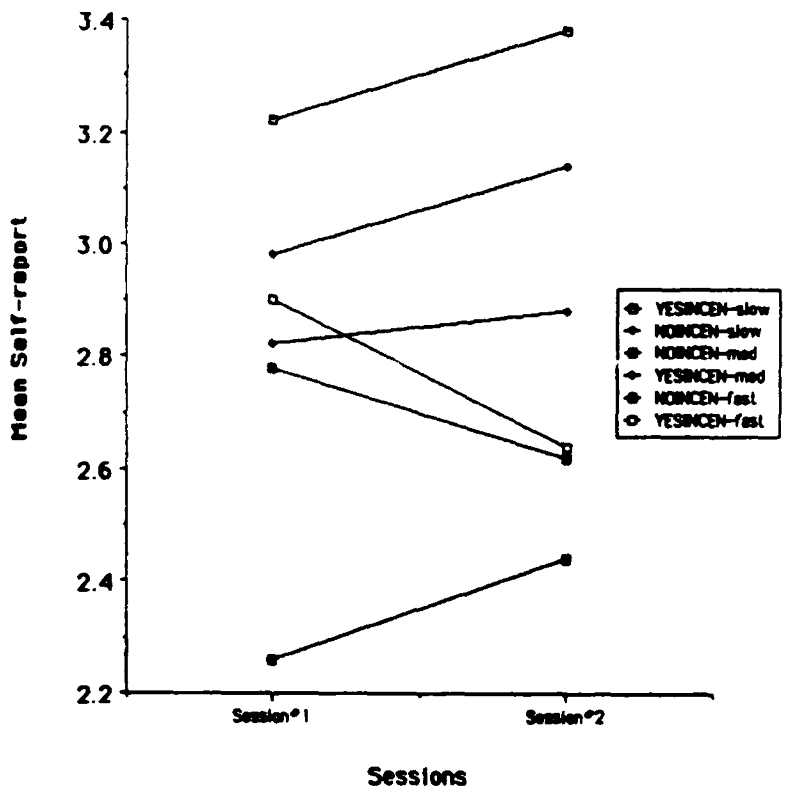

Figure 15. Group x speed x sessions interaction for Question 4: Mental counters.

Summary of Incentive Effects

1. Performance level. MC was the only test on which the subjects performance was significantly boosted by incentives. It is noteworthy that MC was also the most complex task in the study, in that it involves a substantial mental workload.

2. Effort. The questionnaire responses indicated that subjects who received incentives tried harder on all the cognitive speed tests. The physiological indices, however, were apparently not sensitive enough to register the increased activation.

3. Overall. It appears that differences in motivation are not the source of performance differences on simple cognitive tasks such as IT and NIPI, in that level of effort was unrelated to performance on these tasks. The incentive effect for the more complex MC Test, however, raises the possibility that results with Counters are confounded by motivation. In the next section, we explore whether correlations between intelligence and cognitive speed tests (including Counters) are changed by incentives.

4. Correlations Between Cognitive Speed and Intelligence

Tables 14 through 17 present correlational data pertaining to IQ and the relationship between IQ and task performance, and IQ and the physiological measures.

Table 14 shows the intercorrelations among the IQ and GPA variables. The GPAs showed only modest relationships with SAT scores, which is typical for San Diego State University samples (McCornack, 1982). The Advanced Raven, Standard Raven, Advanced Otis-Lennon, and SAT total intercorrelated highly, with a range of .44 to .75. The GPA measures were minimally related to intelligence, however.

Table 14

Intercorrelations Among the IQ and Grade Point Average Variables

Source	ADV.RAV	ST.RAV	SATV	SATQ	SATT	FRGPA	HSGPA	OTIS
ADVRAV	1.00	.65**	.32**	.48**	.48**	.10	.14	.44**
STRAV		1.00	.19	.51**	.44**	.21*	.16	.46**
SATV			1.00	.37**	.80**	.27**	.14	.65**
SATQ				1.00	.86**	.13	.20*	.61**
SATT					1.00	.23*	.20*	.75**
FRGPA						1.00	.25**	.38**
HSGPA							1.00	.13
OTIS								1.00
Mean	22.55	51.39	430.73	496.73	927.47	2.63	3.03	58.10
SD	5.02	4.4	71.64	83.32	128.50	.63	.32	9.37
N	99	71	95	95	95	101	95	98

*P < .05.
**P < .01.

Table 15 shows the correlations among IQ and performance for all subjects at session 1, prior to the incentive manipulation, and at session 2, in which incentive and no incentive groups are combined. Table 15 also includes the correlation of GPA and performance. The variables "COUNTA" and "COUNTB" refer to a MC composite taken at sessions 1 and sessions 2, respectively, which was formed by summing the scores from all three levels of MC. Examination of Table 15 reveals that the performance on the MC Test produced the strongest and most consistent correlations with IQ, a result consistent with those from previous analyses (Saccuzzo & Larson, 1987). The IQ-IT correlations in session 1 were low, but significant for the Advanced Raven, SAT Total, and Advanced Otis, which is also consistent with previous analyses. Of the NIPI variables at session 1, only the standard deviation of the physical identity task (PISD) yielded a significant relationship (-.20 with the Standard Raven and -.18 with the Advanced Otis). With only a few minor exceptions, the IQ-Performance correlations tended to fall slightly in session 2.

Table 15

Correlations Among IQ and Performance:
All Subjects, Sessions 1 and 2

Source	ADV.RAV	ST.RAV	SATV	SATO	SATT	FRGPA	HSGPA	OTIS
First Session								
ITTCA	.20*	.17	.15	.15	.18*	.07	-.18*	.18*
MCTSA	.24**	.31**	.00	.50**	.32**	.04	.11	.22*
MCTMA	.33**	.38**	.14	.23*	.23*	-.02	.08	.22*
MCTEA	.36**	.42**	-.04	.24**	.13	.03	.08	.08
COUNTA	.39**	.44**	.03	.40**	.28**	.02	.11	.21*
PIMEDA	-.03	.03	-.07	.11	.03	-.00	.11	-.04
PISDA	-.20*	-.11	-.13	-.05	-.11	-.10	.00	-.18*
NIMEDA	.10	-.02	.04	.05	.06	-.02	-.00	-.00
NISDA	-.11	-.22	-.03	-.06	-.06	.01	-.17	-.15
Retest Session								
ITTCB	.18*	.09	.09	.08	.10	.16	-.13	.16
MCTSB	.33**	.45**	.13	.33**	.22**	.00	.09	.22*
MCTMB	.19*	.19	.12	.29**	.26**	.13	.12	.24**
MCTEB	.22*	.35**	-.02	.22*	.13	-.06	-.03	.14
COUNTB	.29**	.39**	.08	.32**	.26**	.01	.06	.23**
PIMEDB	.02	-.18	-.04	-.03	-.04	.03	.07	-.16
PISDB	-.17	-.25*	-.04	.01	-.01	.10	-.07	-.13
NIMEDB	.03	-.14	.09	-.03	.03	-.04	-.13	-.13
NISDB	-.09	-.14	-.03	-.08	-.07	-.11	-.13	-.19*

*P < .05.
**P < .01.

Despite the drops in correlations, which might be attributed to practice effects and task automation (Ackerman, 1986), the correlations between MC and IQ remained significant. Performance at session 2 could have been influenced by two variables--incentives and practice. The variables can be untangled by examining the correlations for the incentive and no incentive groups separately for sessions 1 and 2 (see Tables 16 and 17).

Tables 16 and 17 present the IQ-Performance correlations at sessions 1 and 2 for the no incentive and incentive groups, respectively. Both tables include a composite, IQ, which is based on an equal weighing of the Advanced Raven, SAT Total, and Advanced Otis. The Standard Raven was not included due to the relatively few subjects for which scores were available. Inspection of Tables 16 and 17 reveals that MC was the most stable and consistent correlate of IQ. For the IT task, the groups were initially quite different before the incentive manipulation. These differences illustrate the erratic nature of the IT-IQ relationship, as noted in previous research (Saccuzzo, Larson, & Rimland, 1986; Saccuzzo & Larson, 1987). Inspection of Tables 16 and 17 further reveals that the IQ-IT relationship varied widely for the different IQ measures. Similarly, there were obvious initial group differences in the IQ-Performance correlation for PISD, and the strength of the correlation also varied with the IQ measure.

Table 16

Correlations Between IQ and Performance:
No Incentive Group, Sessions 1 and 2

Source	ADVRAV	STANDRAV	OTIS	SATT	IQ
ITTCA	.01	.06	.26*	.23	.20
ITTCB	.09	.04	.27*	.20	.26*
MCTSA	.29*	.33*	.26*	.20	.27*
MCTSB	.35**	.45**	.25*	.20	.28*
MCTMA	.44**	.57**	.27*	.11	.30*
MCTMB	.19	.25	.23	.25*	.26*
MCTFA	.47*	.47**	.02	.09	.22
MCTFB	.30*	.60**	.29*	.18	.28*
COUNTA	.48**	.53**	.21	.15	.30*
COUNTB	.34**	.53**	.30**	.23	.32*
PIMEDA	-.09	-.04	.03	.10	-.01
PIMEDB	-.08	-.32*	-.06	-.03	-.10
PISDA	-.15	-.04	.05	.10	.03
PISDB	-.22	-.37*	-.11	.01	-.19
NIMEDA	.02	.00	.00	.16	.08
NIMEDB	.01	-.23	-.03	.08	.01
NISDA	-.03	-.18	-.09	.02	-.07
NISDB	-.14	-.28	-.19	-.13	-.20

*P < .05.
**P < .01.

Table 17

Correlations Between IQ and Performance:
Incentive Group, Sessions 1 and 2

Source	ADVRAV	STANDRAV	OTIS	SATT	IQ
ITTCA	.33**	.24	.05	.05	.13
ITTCB	.29*	.19	.06	-.01	.11
MCTSA	.17	.23	.15	.49**	.32**
MCTSB	.28*	.45**	.22	.42**	.49**
MCTMA	.17	.02	.15	.41**	.34*
MCTMB	.17	.13	.26*	.27*	.31*
MCTFA	.27*	.33*	.12	.17	.34*
MCTFB	.09	-.00	-.01	.09	.07
COUNTA	.27*	.26	.17	.44**	.43**
COUNTB	.21	.23	.17	.32*	.35*
PIMEDA	-.15	-.01	-.16	-.14	-.24
PIMEDB	.10	-.03	-.25*	-.06	-.01
PISDA	-.29*	-.17	-.35**	-.28*	-.41**
PSIDB	-.11	-.12	-.16	-.05	-.24
NIMEDA	.15	-.05	.01	-.06	.02
NIMEDB	.04	-.03	-.25*	-.01	-.07
NISDA	-.21	-.20	-.16	-.12	-.27
NISDB	-.03	.02	-.19	-.01	-.12

*P < .05.
**P < .01.

We noted the relative stability of results with MC, compared to IT and PISD. Closer inspection of the correlational results for MC reveals that neither incentives nor practice had much effect on the IQ-Performance correlation. For the no incentive group, the Performance-IQ composite correlation was .30 (p < .05) at session 1 and 0.32 (p < .05) at session 2, showing minimal change with practice. For the incentive group, the correlation dropped slightly, but nonsignificantly (Z = .6836), as revealed by the Z test (Glass & Stanley, 1979). Thus, neither incentives nor practice had an appreciable effect on the IQ-Performance relationship. The within-group validity differences between sessions 1 and 2 for most of the tasks were less than the validity difference between the incentive and no incentive groups at session 1, prior to any manipulation. If anything, there was a regression to the mean in both groups.

DISCUSSION

The results revealed that incentives had no effect on performance on the IT and NIPI tests. Whereas, subjects in the incentive group reported that they either tried harder or used more effort at session 2 than the no incentive subjects, there was no corresponding increase in performance. Apparently, there was a limit to performance which further effort could not surmount.

Analyses of performance for the MC Test showed a somewhat different pattern than found for IT and NIPI. A significant Group X Sessions interaction was found, which indicated that while both groups improved with practice, the incentive group improved to a greater degree. For the MC Test, incentives led to better performance. Thus, in contrast to performance on the relatively simple reaction time and IT tests, the MC Test was susceptible to incentives. Apparently, then, performance on the version of the MC Test used in the present study is less "hard wired" and more subject to strategic control and resource allocation than IT and NIPI. The Group X Sessions effect for MC was paralleled by a Group X Sessions effect for questions 1 and 3 of the self-report questionnaire. For both questions, the interaction effects indicated that subjects in the incentive group said that they tried harder (question 1) or had less reserve effort (question 3). However, the same two interactions were found with questionnaire response to IT with no corresponding increases in performance. Thus, while trying harder or using more effort led to increases on the MC Test, it did not lead to similar increases for IT.

In further support of a relationship between task difficulty and incentive effects is the Group X Speed X Sessions triple interaction for MC (see Figure 7). This interaction revealed that rate of improvement for the incentive group varied as a function of speed (task difficulty). The faster the speed, the greater the effect of incentives. Thus, when the task is relatively easy, incentives have little effect, because there is little room for improvement. As task difficulty increases, however, there is greater room for improvement when more effort is expended.

In sum, in answer to the question, "How does motivation, induced through incentives, affect task performance on cognitive speed tests?," the answer appears to be as follows. In the typical research setting, for IT and simple reaction time tasks such as NIPI, motivation has little effect on performance. While subjects in the incentive group said they tried harder on session 2 for IT and NIPI, compared to the no incentive group (see Figures 9 and 11), there was no corresponding increase in performance. For the more difficult and complex MC Test, by contrast, incentives do lead to performance increases. These conclusions pertaining to incentives are, of course, limited to the range of motivation in a typical research setting. The picture could be quite different for completely unmotivated or antagonistic subjects.

Independent of the question of how motivation affects performance is the question of how motivating conditions affect the IQ-Performance correlation. This latter question was addressed primarily in Tables 15 through 17. The correlations for the incentive and no incentive groups for sessions 1 and 2 between the MC composite and the IQ composite reflect the clear trend of the data. Essentially, there was little change in the IQ-Performance correlation whether or not subjects had incentives. The differences between the groups prior to the incentive manipulation was almost twice the difference found between sessions 1 and 2 for the incentive group. Thus, although incentives led to increased performance on the MC Test, they did not affect significantly the IQ-Performance correlation. The correlation between IQ and performance for motivated subjects is just as strong as it is for random groups. Clearly, then, the IQ-MC correlation found here and in previous studies cannot be attributed to motivation alone. Some other process such as some basic underlying aspect of intelligence is needed to account for the well documented IQ-Performance correlation. Results for IT and the NIPI test are less clear, and within-group validities again differed prior to the introduction of incentives. In general, however, the overall results of the study do not show that motivation plays a particularly important role in results with the IT and NIPI paradigms.

The present study failed to find a clear relationship between arousal and performance, arousal and incentives, and arousal and IQ. One problem in this regard may have been the relative simplicity of the tasks, most notably, NIPI. As long as resources keep up with demands, a task need not be arousing. Thus, in some cases the tasks were well within the capacity of subjects. When task difficulty clearly exceeded capacity, performance broke down. It was only when task difficulty increased between MC slow and medium, that a significant physiological effect (i.e., change in SC) was found. Interestingly, SC levels were significantly higher at MC medium than at MC hard, where performance deteriorated. From these findings we might conclude that when task difficulty increases only to the degree where increased effort can lead to better performance (i.e., when there is room for improvement given capacity limitations), the increased difficulty is arousing and reflects greater expenditure of effort. If, however, the increase in difficulty exceeds capacity, the task is less arousing and performance deteriorates. Nevertheless, arousal differences between subjects and changes in physiological arousal, as measured here, reveal nothing about individual differences in intelligence.

In addition to elucidating the role of motivation and performance for the three types of tasks studied herein and showing that certain IQ-Performance correlations are not, in fact, attributable to motivation, present results lead to a number of conclusions of significance to the Armed Forces. First, mental speed, as indexed by IT and reaction time, appears to be hard wired, that is, relatively insensitive to motivation and incentives but highly vulnerable to difficulty manipulations (either faster speed or greater complexity). These tasks, however, are of limited predictive value in that the IQ-Performance correlation is relatively weak and varies widely across samples.

The MC Test, by contrast, is extremely promising. First, it clearly distinguishes subjects of varying ability and has the highest and most stable correlations with IQ. Second, the MC Test is extremely flexible. Task difficulty can be varied both in terms of the number of targets and in terms of the speed of target presentation. Hence, it is possible to devise a large set of items of increasing difficulty (by varying speed) and complexity (by varying the number of targets). The MC Test can be further varied such that greater increases in difficulty do not necessarily lead to a breakdown in performance, as was revealed in the present study. Thus, it is possible to use MC to evaluate strategic control and resource deployment. Given its flexibility, the MC Test appears to have considerable potential in adding incremental validity to existing test procedures.

CONCLUSIONS

1. Incentives have no appreciable effect on performance for IT and relatively simple reaction time tasks; but do affect performance on more complex or difficult tests such as the more difficult level of MC.

2. Whether or not subjects are extrinsically motivated, the IQ-Performance relationship remains about the same. Motivation alone cannot explain the IQ-Performance relationship found for cognitive speed tests.

3. There is little or no relationship between performance on cognitive speed tests and physiological arousal except when increases in task difficulty are small enough to permit constant performance through increased effort or resource allocation.

4. There is little or no relationship between intelligence and arousal, as indexed by HR and SC changes.

5. Changes in self-reported effort are not necessarily correlated with changes in performance or physiological arousal.

6. The IT and NIPI tests are of questionable value for personnel selection due to widely varying predictive validity coefficients.

7. The MC Test has considerable potential for adding incremental validity to existing batteries.

RECOMMENDATIONS

Additional investigation of the properties of the MC Test is warranted. Further consideration of the validity and ideal complexity/difficulty variations may be desirable as a prelude to using the MC Test on an extensive scale as an aid to the selection process and in adding incremental validity to existing test batteries.

REFERENCES

Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. Intelligence, 10, 101-139.

Barrett, P., Eysenck, H. J., & Lucking, S. (1986). Reaction time and intelligence: A replicated study. Intelligence, 10, 9-40.

Brand, C. R., & Deary, I. J. (1982). Intelligence and "inspection time." In H. J. Eysenck, (Ed.), A model for intelligence. New York: Springer-Verlag.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Felsten, G., & Wasserman, G. S. (1980). Visual masking: Mechanisms and theories. Psychological Bulletin, 88, 329-353.

Geiselman, R. E., Woodward, J. A., Beatty, J. (1982). Individual differences in verbal memory performance: A test of alternative information processing models. Journal of Experimental Psychology: General, 111, 109-134.

Glass, G. V., & Stanley, J. C. (1970). Statistical methods in education and psychology. New Jersey: Prentice-Hall.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance. Vol. II: Cognitive processes and performance. New York: John Wiley.

Jensen, A. R. (1982). Reaction time and psychometric "g." In H. J. Eysenck (Ed.), A model for intelligence. New York: Springer-Verlag.

Jensen, A. R. (1987a). Process differences and individual differences in some cognitive tasks. Intelligence, 11, 107-136.

Jensen, A. R. (1987b). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), Speed of information processing and intelligence. Norwood, NJ: Ablex.

Jensen, A. R. (1987c). The g beyond factor analysis. In J. C. Conoley, J. A. Glover, & R. R. Ronning (Eds.), The influence of cognitive psychology on testing and measurement. Hillsdale, NJ: Lawrence Erlbaum.

Kahneman, D. (1973). Attention and effort. New Jersey: Prentice-Hall.

Keating, D. P., & MacLean, D. J. (1987). Cognitive processing, cognitive ability, and development: A reconsideration. In P. A. Vernon (Ed.), Speed of information processing and intelligence. Norwood, NJ: Ablex.

Lacey, B. C., & Lacey, J. I. (1974). Studies of heart rate and other bodily processes in sensorimotor behavior. In P. A. Obrist, A. H. Black, J. Brenner, & L. V. DiCara (Eds.), Cardiovascular psychophysiology. Chicago: Aldine, 526-564.

Larson, G. E. (1986). The Mental Counters Test. Submission to the Joint-Service Future Testing Committee.

Larson, G. E., Merritt, C. R., & Lattin, K. E. (1988). Reliability and construct validity of reaction time, inspection time, and machine-paced tests of cognitive speed (NPRDC Tech. Note 88-37). San Diego: Navy Personnel Research and Development Center.

Larson, G. E., & Rimland, B. (1984). Cognitive speed and performance in basic electricity and electronics (BE&E) school (NPRDC Tech. Rep. 85-3). San Diego: Navy Personnel Research and Development Center.

Lawler, K. A., Obrist, P. A., & Lawler, J. E. (1976). Cardiac and somatic response patterns during a reaction time task with children and adults. Psychophysiology, 13, 448-455.

Marr, D. B., & Sternberg, R. J. (1987). The role of mental speed in intelligence: A triarchic perspective. In P. A. Vernon (Ed.), Speed of information processing and intelligence. Norwood, NJ: Ablex, 271-294.

McCornack, R. L. (1982). Predicting freshman grades and the analysis of prediction errors in four groups of minority students. San Diego: Institutional Research, San Diego State University.

Nettlebeck, T. (1982). Inspection time: An index for intelligence? Quarterly Journal of Experimental Psychology, 34A, 299-312.

Nettlebeck, T., & Kirby, N. H. (1983). Measures of timed performance and intelligence. Intelligence, 7, 39-52.

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. Psychological Review, 74, 392-409.

Saccuzzo, D. P., & Larson, G. E. (1987). Analysis of test-retest reliability for a battery of cognitive speed tests (NPRDC Tech. Rep. 88-10). San Diego: Navy Personnel Research and Development Center.

Saccuzzo, D. P., Larson, G. E., & Rimland, B. (1986). Speed of information processing and individual differences in intelligence (NPRDC Tech. Rep. 86-23). San Diego: Navy Personnel Research and Development Center.

Vernon, P. A. (1983). Speed of information processing and general intelligence. Intelligence, 7, 53-70.

Vernon, P. A., Nador, S., & Kantor, L. (1985). Reaction times and speed-of-processing: Their relationship to timed and untimed measures of intelligence. Intelligence, 9, 357-374.

Weinstein, L. (1981). Incentive contrast effects in humans with monetary reinforcement and reaction time. Acta Psychologica, 47, 83-87.

Winer, B. J. (1962). Statistical principles in experimental design. New York: McGraw-Hill.

APPENDIX A

SELF-REPORT QUESTIONNAIRE

APPENDIX A

Self-Report Questionnaire

Subject Number _______

Name _______________

Session ___1____or___2

1. How hard did you try on this task?
(circle one)

6	5	4	3	2	1
Extremely Hard	Very Hard	Hard	Somewhat hard	Hardly	Not at all.

2. How difficult was this task for you?

6	5	4	3	2	1
Extremely Difficult	Very Difficult	Difficult	Somewhat Difficult	Very Easy	Extremely Easy

3. How much better do you think you could have done, had you used more effort? (Circle one)

6	5	4	3	2	1
Extremely Better	Very Much Better	Much Better	Somewhat Better	Slightly Better	Not better At All

4. How much more effort could you have expended, had the task been more difficult? (Circle one)

6	5	4	3	2	1
Very Much More	Considerable More	Much More	Somewhat More	Slightly More	No More

1

DISTRIBUTION LIST

Distribution:
Assistant Secretary of Defense (Force Management and Personnel)
Assistant for Manpower Personnel and Training Research and Development (OP-01B2)
Technology Area Manager, Office of Naval Technology (Code 222)
Defense Technical Information Center (DTIC) (2)

Copy to:
Director, Office of Naval Research (OCNR-10)
Head, Cognitive and Neuro Sciences (OCNR-1142)
Perceptual Science (OCNR-1142PS)
Cognitive and Decision Science (OCNR-1142CS)
Office of Naval Research, London
Commanding Officer, Naval Aerospace Medical Research Laboratory, Pensacola, FL
Technical Director, U.S. ARI, Behavioral and Social Sciences, Alexandria, VA (PERI-ZT)
Chief, U.S. ARI-USAREUR (Library) (2)
Manpower and Personnel Division (AFHRL/MO)
Scientific and Technical Information (STINFO) Office
TSRL/Technical Library (FL 2870)
Commanding Officer, U.S. Coast Guard Research and Development Center, Avery Point, Groton, CT
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Institute for Defense Analyses, Science and Technology Division

Courtesy copies:
Dr. L. Valentine, AFHRL
Dr. A. Divgi, CNA
Dr. W. Tirre, AFHRL
Dr. F. Schmidt, U. of Iowa
Dr. B. Green, Johns Hopkins
Dr. S. Embretson, U. of Kansas
Dr. B. Bloxum, DMDC, Monterey